Wenceslao J. González (coord.)

Diversidad de la explicación científica



DIVERSIDAD DE LA EXPLICACIÓN CIENTÍFICA

Ariel Filosofía

Wenceslao J. González

DIVERSIDAD DE LA EXPLICACIÓN CIENTÍFICA

Ariel



Diseño de la cubierta: Joan Batallé

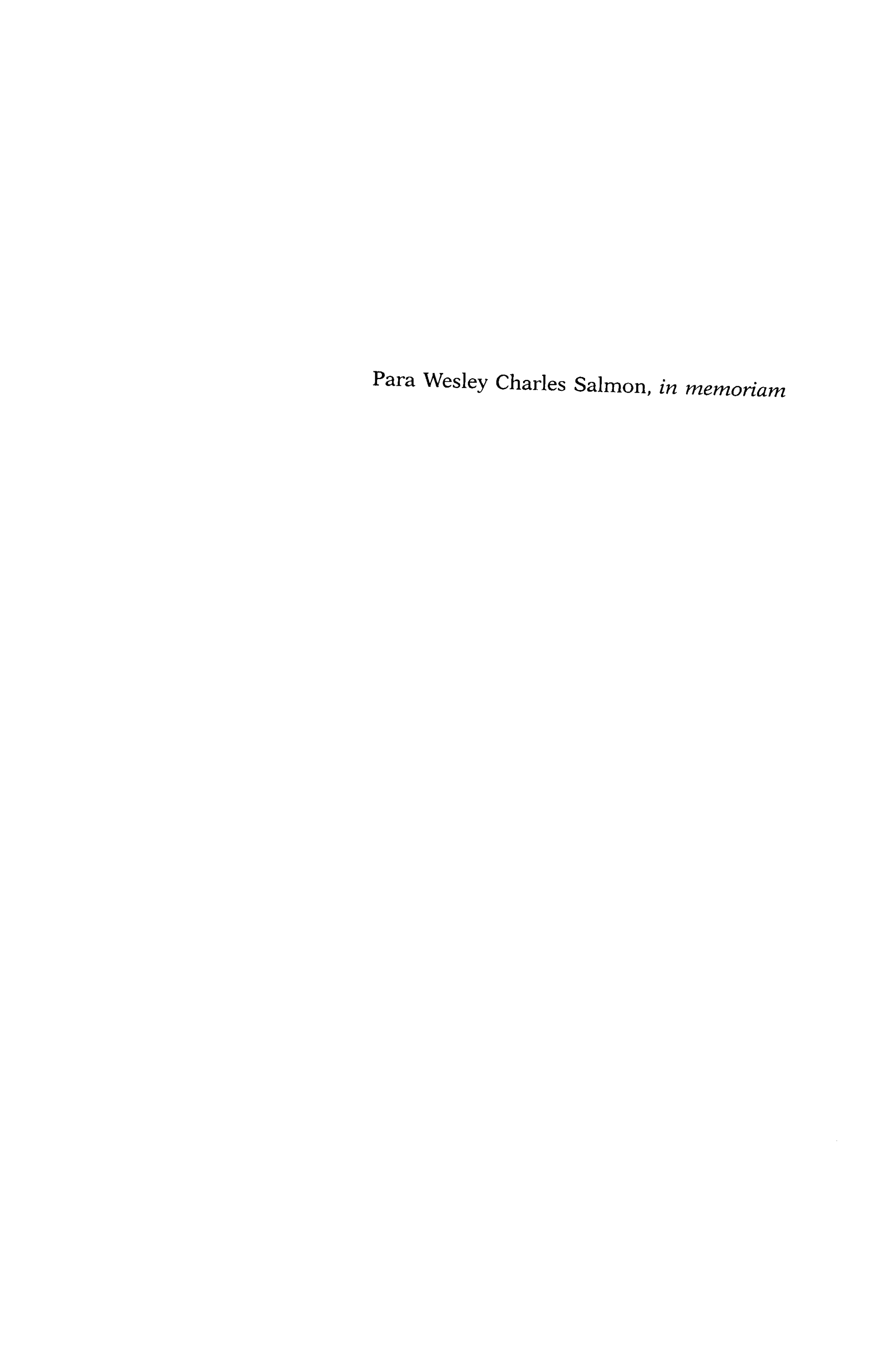
1.ª edición: junio 2002

ISBN: 84-344-8763-2

Depósito legal: B. 24.225 - 2002

Impreso en España

2002. – A&M Gràfic, S. L.
Polígono Industrial «La Florida»
08130 Santa Perpètua de Mogoda
(Barcelona)



Para Wesley Charles Salmon, *in memoriam*

ÍNDICE GENERAL

PARTE I

MARCO TEÓRICO

PARTE II

EVOLUCIÓN DEL PROBLEMA DE LA EXPLICACIÓN CIENTÍFICA

PARTE III

PROPUESTAS PARA LAS EXPLICACIONES CIENTÍFICAS

PARTE IV

LAS EXPLICACIONES CAUSALES: ESTRUCTURA Y APLICACIÓN

PARTE V

EXPLICACIONES TELEOLÓGICO-CAUSALES Y EXPLICACIONES PROGRAMÁTICAS

PARTE I

MARCO TEÓRICO

CAPÍTULO 1

CARACTERIZACIÓN DE LA «EXPLICACIÓN CIENTÍFICA» Y TIPOS DE EXPLICACIONES CIENTÍFICAS

WENCESLAO J. GONZÁLEZ

1. Dos planos ante las variedades de explicación científica

Plantear la cuestión de cuáles son las variedades de explicación científica puede llevar, al menos inicialmente, a dos planos distintos: en primer lugar, a las variaciones sobre cómo concebir la «explicación científica» en cuanto tal —su estatuto y características—; y, en segundo término, a las diversas posibilidades de explicar en la Ciencia a tenor del contenido (esto es, los tipos de explicaciones en razón del rasgo resaltado). Esto requiere atender a la práctica de los científicos —cómo entienden la explicación científica y qué tipo de explicaciones aceptan—; pero ha de contemplar también lo que se podría llegar a desarrollar para incrementar el avance científico.

Esa polaridad temática de los dos planos —la aclaración de lo definitorio de la «explicación científica» y la indagación acerca de los tipos de explicaciones aceptables— late en el presente volumen. Cada uno de esos aspectos ha recibido la atención de Wesley Charles Salmon, que ha analizado en detalle muchos de los problemas que se han suscitado durante décadas. De hecho, en su último libro como autor reconocía «haber estado trabajando activamente sobre la explicación científica durante más de treinta años».[1] Por eso, tanto este trabajo como otros posteriores del presente volumen le conceden una especial relevancia.

Acerca del primer plano —las variaciones en cuanto a cómo entender la explicación científica misma— hay, en efecto, observacio-

1. SALMON, W. C., *Causality and Explanation*, Oxford University Press, Nueva York, 1998, p. ix.

nes a lo largo del célebre libro de W. C. Salmon *Four Decades of Scientific Explanation.*[2] Es un volumen que ha servido de inspiración para la presente publicación. Porque, una vez abordadas ahí las cuatro primeras décadas de la evolución de la explicación científica (1948-1987), hacía falta atender al periodo siguiente, que es el cometido de buena parte de los trabajos aquí reunidos. También en *Explaining Explanation,*[3] de David H. Ruben, se ofrecen reflexiones sobre el problema —qué es una «explicación»—: su enfoque es más amplio, en cuanto que atiende en detalle a importantes filósofos del pasado, tanto clásicos como del siglo XIX. Después analiza posturas de la segunda mitad del siglo XX, cuando la preocupación central por la «explicación científica» cobra protagonismo.

Sobre el segundo plano —los tipos de explicaciones científicas desde el punto de vista del contenido— han trabajado un buen número de autores. Entre ellos está Ernest Nagel, que marcó una época con su influyente estudio *The Structure of Science,*[4] que tiene un subtítulo muy expresivo: «Problemas de la Lógica de la explicación científica». El libro ofrece una interesante distinción entre cuatro tipos de explicaciones científicas: el modelo deductivo, las explicaciones probabilísticas, las explicaciones funcionales o teleológicas y las explicaciones genéticas.[5] Esas cuatro variedades son particularmente relevantes cuando se ven desde la perspectiva histórica: ponen de relieve que, entre los autores del Empirismo lógico —herederos del ideal de la Ciencia unificada del Neopositivismo—, se acepta expresamente que la Ciencia no se desarrolla mediante un único tipo de explicaciones sino a través de una *diversidad* de explicaciones. Una de ellas es la explicación causal, que ha sido durante años uno de los ejes de coordenadas del Pensamiento de Salmon.[6]

Si se acepta que la «explicación científica» es un concepto metodológico, de modo que mira al avance riguroso del conocimiento,

2. SALMON, W. C., *Four Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis, 1990.

3. RUBEN, D. H., *Explaining Explanation*, Routledge, Londres, 1990. También reviste especial interés su compilación RUBEN, D. H. (ed.), *Explanation*, Oxford University Press, Oxford, 1993.

4. NAGEL, E., *The Structure of Science. Problems in the Logic of Scientific Explanation*, Harcourt, Brace and World, N. York, 1961.

5. Cfr. NAGEL, E., *The Structure of Science*, pp. 20-26.

6. Su libro *Causality and Explanation* ofrece 26 trabajos sobre este tema, de los cuales 7 no habían sido publicados antes de 1998. Su preocupación por la explicación causal alcanzó hasta sus últimos días y queda plasmada en la interesante visión sistemática ofrecida en SALMON, W. C., «Estructura de la explicación causal», en GONZÁLEZ, W. J. (ed), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 141-159.

y si se asume que lo propio de la Metodología de la Ciencia es atender tanto al *ser* —su realidad como actividad humana— como al *deber ser* de la Ciencia —la búsqueda de los mejores métodos posibles—,[7] entonces hay que examinar una doble vertiente: cómo explican los científicos al desarrollar su actividad y cómo pueden mejorar los procesos seguidos de modo que se dé progreso científico.[8] Ambos aspectos atañen a los dos planos antes señalados: la caracterización de la «explicación científica» en cuanto tal y la modulación de sus diversas variantes desde el punto de vista del contenido.

2. El problema de la caracterización de la «explicación científica»

Desde el punto de vista de la caracterización de la «explicación científica», Salmon señala que hay una distinción fundamental entre saber algo (conocimiento *that*) y saber por qué (conocimiento *why*): «una cosa es saber *que* cada planeta periódicamente invierte (*reverses*) la dirección de su movimiento con respecto al trasfondo (*background*) de las estrellas fijas; y otra, muy distinta, es saber *por qué*. El conocimiento del primer tipo es descriptivo; el conocimiento del segundo tipo es explicativo. Lo que proporciona la comprensión científica del mundo es el conocimiento explicativo (*explanatory knowledge*)».[9] La explicación es entonces, para Salmon, un *contenido cognitivo* que da respuesta a un interrogante concreto y lo hace de modo acorde con el resto del conocimiento poseído o visión científica de la realidad.

Tal como concibe esta noción en la Ciencia, parece claro que Salmon remite a los *hechos* —sean sucesos singulares o tendencias generales— del mundo, es decir, no mira primariamente a la *actividad* científica en cuanto tal (esto es, la tarea desarrollada por los científicos). Explicar es entonces «dar razón de lo que sucede», una vez que se conoce qué sucede en la realidad estudiada. En otras palabras, caracteriza la explicación científica como «un intento de hacer comprensible o inteligible algún suceso particular (como el ac-

7. Cfr. GONZÁLEZ, W. J., «Ámbito y características de la Filosofía y Metodología de la Ciencia», en GONZÁLEZ, W. J. (ed.), *Aspectos metodológicos de la investigación científica*, 2ª edición, Ediciones Universidad Autónoma de Madrid y Publicaciones de la Universidad de Murcia, Madrid-Murcia, 1990, pp. 49-78; en especial, pp. 72-73.

8. GONZÁLEZ, W. J., «Progreso científico, Autonomía de la Ciencia y Realismo», *Arbor*, v. 135, n. 532, (1990), pp. 91-109; y GONZÁLEZ, W. J., «Progreso científico e innovación tecnológica: La 'Tecnociencia' y el problema de las relaciones entre Filosofía de la Ciencia y Filosofía de la Tecnología», *Arbor*, v. 157, n. 620, (1997), pp. 261-283.

9. SALMON, W. C., *Four Decades of Scientific Explanation*, p. 3.

cidente de 1986 en la Central nuclear de Chernobyl) o algún hecho general (como el color cobre de la Luna durante un eclipse total) mediante la atención a otros hechos, concretos y/o generales, obtenidos a partir de una o más ramas de la Ciencia Empírica».[10]

2.1. EXPLICACIONES EN LA CIENCIA Y «EXPLICACIONES CIENTÍFICAS»

Advierte Salmon que el término «explicación» se emplea con frecuencia de modo que poco o nada tiene que ver con la *explicación científica.*[11] Recuerda que Michael Scriven se quejaba de uno de los modelos de explicación científica propuestos por Carl Gustav Hempel, en cuanto que no encajaría siquiera con su uso ordinario, como en el caso de alguien que explica, mediante gestos, a un mecánico yugoslavo que no sabe inglés qué es lo que no funciona en su coche. Ante esa objeción, considera que Hempel dio una respuesta correcta:[12] esa posición sería comparable a quejarse de la explicación precisa que se da a la palabra «prueba» (*proof*) en Matemáticas, cuestionándola porque no recoge el significado del término «prueba» presente en afirmaciones de la vida ordinaria como 'la prueba de una tarta está en comerla'.[13]

Aunque, para Salmon, la explicación científica comporta responder a la pregunta «por qué sucede algo» (*why something occurs*), admite que en la Ciencia están también presentes *otros tipos de explicaciones*, tales como las explicaciones semánticas (*what something means*) y las explicaciones sobre acciones, esto es, cómo llevar a cabo determinadas actuaciones (*how to perform certain activities* o *how to do something*). Así, señala que «las explicaciones de significados se dan de hecho en las Ciencias; para encontrar el significado de un término técnico, se puede consultar un manual científico. Las explicaciones sobre cómo realizar diversas actividades también se encuentran en la Ciencia: un científico puede explicar a un técnico cómo construir un tipo particular de aparato de detección».[14] Pero, a continuación, insiste en que ése no es el caso ordi-

10. SALMON, W. C., «Scientific Explanation», en SALMON, M. ET AL., *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs, NJ, 1992, p. 8.

11. Cfr. SALMON, W. C., *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, N. J., 1984, pp. 9-11.

12. SALMON, W. C., *Four Decades of Scientific Explanation*, p. 6.

13. Cfr. HEMPEL, C. G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, N. York, 1965, p. 6.

14. SALMON, W. C., *Causality and Explanation*, p. 6.

nario cuando se habla de «explicaciones científicas». Considera, en efecto, que lo habitual en la Ciencia son las explicaciones que preguntan *por qué* ocurre un fenómeno concreto, sea éste un hecho particular o una regularidad de carácter general.

Sin embargo, no toda pregunta *por qué* remite a una explicación científica: también se puede preguntar *por qué* en casos de búsqueda de justificación, de consuelo o de pruebas empíricas (*evidence*). El primer caso se da cuando una mujer inquiere por qué recibe un salario inferior a un colega masculino que hace el mismo trabajo; el segundo supuesto se aprecia cuando unos padres se interrogan por la pérdida repentina de un hijo; y el tercero se ve al preguntar «¿por qué hemos de creer que las galaxias distantes se alejan de nosotros a gran velocidad?». A diferencia de estos casos, Salmon piensa que la pregunta «por qué» orientada a la búsqueda de la explicación suele ser: ¿por qué sucede? (*why-does*).[15] Paralelamente deja abierta la posibilidad a que algunas explicaciones no puedan ser cubiertas atendiendo a las preguntas «por qué», de modo que puedan requerir otras preguntas. Podría ser el caso de las cuestiones sobre cómo es posible algo (*how-possibly-questions*) o los interrogantes acerca de cómo ha sido realmente algo (*how-actually form*).[16]

Conviene insistir en que esto comporta una importante variación respecto de modos anteriores de entender la «explicación científica». Porque ahora ya no se ve como *argumento*, como hacía Carl Gustav Hempel[17] (y también Paul Oppenheim); ni es tampoco un mero *enunciado* —como sí puede ser una predicción o, en su caso, una confirmación—: es la respuesta a interrogantes de una determinada índole. Entre otros autores, ha sido Bas C. van Fraassen quien ha resaltado la superación de esas posiciones anteriores sobre la explicación científica. Lo ha hecho cuando ha propuesto una pragmática de la explicación:[18] «una explicación no es lo mismo que una proposición, o un argumento, o una lista de proposiciones: es una *respuesta*. (De modo análogo, un hijo no es lo mismo que un hombre, incluso si todos los hijos son hombres y todo hombre es hijo). Una explicación es una respuesta a una

15. Cfr. *Causality and Explanation*, p. 6.

16. Cfr. SALMON, W. C., «Scientific Explanation», pp. 9-10.

17. Cfr. HEMPEL, C. G., *Philosophy of Natural Science*, Prentice Hall, Englewood Cliffs, NJ, 1966, p. 50.

18. Señalar esto no quiere decir que Bas C. van Fraassen acierte cuando propone dejar al margen de la explicación científica las cuestiones que no sean pragmáticas. Una crítica al respecto se encuentra en KITCHER, PH., «Explanatory Unification», *Philosophy of Science*, v. 48, (1981), pp. 507-531.

pregunta por qué. Así, una teoría de la explicación debe ser una teoría de las preguntas por qué (*why-questions*)».[19]

2.2. Explicación y predicción: De la simetría a la asimetría

Entendida la explicación en los términos aquí expuestos, parece difícil que quepa una simetría estructural entre «explicación» y «predicción», como la propugnada en el célebre artículo de C. G. Hempel y P. Oppenheim.[20] Fue una posición lógico-metodológica que tuvo singular resonancia, pues influyó en conocidos filósofos de la Ciencia, entre los que destaca Adolf Grünbaum. Según su análisis, la diferencia entre ambas —explicación y predicción— sería de índole temporal: dependería de si el suceso *s* es anterior o posterior al «ahora» de la persona que infiere *s* a partir de una serie de condiciones en conjunción con leyes adecuadas. En tal caso, no habría una auténtica divergencia estructural, esto es, no se daría una separación metodológica genuina entre explicar y predecir.[21] Pero esto es algo que rechaza Salmon. Es una posición que cuestiona de una manera realmente respetuosa, pues tiene claro que es una tesis largamente defendida por uno de sus mejores amigos. Piensa, en efecto, que la alegada anisotropía temporal de la explicación científica no permite llegar a una simetría estructural entre explicación y predicción.[22]

Ante este problema planteado por Grünbaum, Salmon va a las raíces y examina las dos partes de la *tesis de la simetría* de Hempel: i) dentro del marco de las circunstancias adecuadas, cada explicación científica puede servir como predicción (esto es, dar un argumento que explique un suceso del pasado —un eclipse de sol, por ejemplo— puede servir para predecir el evento en el futuro); y ii)

19. Fraassen, B. C. van, «The Pragmatic Theory of Explanation», en Pitt, J. C. (ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988, pp. 137-138.

20. Cfr. Hempel, C. G. y Oppenheim, P., «Studies in the Logic of Explanation», *Philosophy of Science*, v. 15, (1948), pp. 135-175; compilado en Hempel, C. G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, pp. 245-295.

21. Cfr. Grünbaum, A., «Temporally Asymetric Principles, Parity between Explanation and Prediction, and Mechanism versus Teleology», *Philosophy of Science*, v. 29, (1962), pp. 146-170. En conversaciones mantenidas en Pittsburgh en el año 1993, Grünbaum me expuso en detalle su defensa de la tesis de la simetría entre explicación y predicción: seguía considerando esa posición como correcta.

22. Cfr. Salmon, W. C., «On the Alleged Temporal Anisotropy of Explanation», en Earman, J., Janis, A., Massey, G. y Rescher, N. (eds.), *Philosophical Problems of the Internal and External Worlds*, University of Pittsburgh Press, Pittsburgh, 1993, pp. 229-248.

en las circunstancias adecuadas, toda predicción científica legítima puede servir como explicación científica. Esto segundo es más problemático; lo utiliza para cuestionar la propuesta hempeliana: Salmon acepta la idea de Israel Scheffler de la predicción como *enunciado* acerca del futuro,[23] que contrapone a la concepción de Hempel de la explicación como *argumento*, de modo que la predicción tiene características propias: no puede ser una explicación. En todo caso, lo más que se podría mantener es que las predicciones científicas legítimas constituyen las *conclusiones* de argumentos que se conforman a unos esquemas concretos de explicación (como el nomológico-deductivo o el probabilístico inductivo), posición que Salmon considera que corresponde a Grünbaum.[24]

Pero, atendiendo a los fenómenos de las Ciencias de la Naturaleza, Salmon entiende que el planteamiento de la simetría carece de base: «la *alegada* anisotropía temporal de la explicación científica es también la anisotropía temporal *real* (*actual*) de la explicación científica».[25] En otras palabras, hay tres anisotropías que van juntas: el tiempo, la causalidad y la explicación científica; y esto es compatible con el hecho de tener leyes de la Naturaleza que tengan simetría temporal (como sucede en la Mecánica newtoniana). Porque la «anisotropía temporal de la explicación científica nada tiene que ver con la reversibilidad o irreversibilidad de los procesos físicos involucrados en la situación. Que haya o que pudiera haber habido un sistema solar justo por encima del nuestro, pero con rotaciones en sentido inverso, es —a mi juicio— irrelevante para el hecho de tener un eclipse de sol *en nuestro sistema solar* en 1919, que era el resultado de las condiciones de su pasado, no de condiciones de su futuro».[26] En suma, como han venido manteniendo diversos autores, entre los que cabe destacar a Nicholas Rescher, hay una asimetría entre la tarea científica de explicar y el quehacer científico de predecir.[27]

Cuál de las dos —explicación o predicción— tiene más peso específico desde el punto de vista metodológico es una cuestión que suscita controversia. La divergencia se constata a lo largo de la se-

23. Cfr. SCHEFFLER, I., «Explanation, Prediction, and Abstraction», *British Journal for the Philosophy of Science*, v. 7, (1957), pp. 293-309.

24. Cfr. SALMON, W. C., «On the Alleged Temporal Anisotropy of Explanation», pp. 231-232.

25. «On the Alleged Temporal Anisotropy of Explanation», p. 245.

26. SALMON, W. C., «On the Alleged Temporal Anisotropy of Explanation», p. 244.

27. Cfr. RESCHER, N., *Predicting the Future*, State University Press New York, N. York, 1998. A este tema dedica buena parte del capítulo 9; en especial, pp. 165-167.

gunda mitad del siglo XX, que es cuando —según Salmon— se supera el nivel científico de la descripción y se considera plausible la «explicación científica». Antes de ese periodo cabe resaltar una tradición predictivista, que cuenta con autores muy influyentes, pertenecientes a tendencias filosóficas diferentes, como F. Bacon, G. W. Leibniz y W. Whewell. Después, durante el periodo que corresponde a las cuatro décadas de explicación científica que examina, resaltan la predicción tanto el falsacionismo popperiano como la Metodología de Programas de Investigación Científica. Esta última concede una clara relevancia metodológica a la predicción: considera que el conocimiento predictivo posee una mayor valoración cognitiva que el explicativo. En Imre Lakatos esto se plasma en una interesante propuesta acerca de un nexo entre predicción científica y hechos nuevos.[28]

Paralelamente, hay otros autores que, en sintonía con tesis empiristas o bien en consonancia con una interpretación bayesiana de la probabilidad de signo personalista o subjetivista, rechazan que la predicción tenga más peso que la explicación científica.[29] También se incide en la Historia de la Ciencia para defender que, con carácter general, la predicción no tiene mayor relevancia cognitiva que la explicación.[30] Aceptar esto supondría admitir que, en conjunto, puede tener tanto valor la acomodación a los hechos por parte de la explicación científica como la predicción de hechos nuevos, bien sea en sentido temporal o en alguna de las otras cinco acepciones posibles de «hechos nuevos».[31]

Aun cuando este tema ha sido analizado con frecuencia, no hay por ahora consenso sobre si una predicción con éxito tiene un carácter confirmatorio mayor que el poseído por la acomodación a los hechos realizada por la explicación. La propuesta de la superioridad

28. Cfr. GONZÁLEZ, W. J., «Lakatos's Approach on Prediction and Novel Facts», *Theoria*, v. 16, n. 42, (2001), pp. 499-518.

29. Cfr. ACHINSTEIN, P., «Explanation v. Prediction: Which Carries More Weight?», en HULL, D., FORBES, M. y BURIAN, R. M. (eds.), *PSA 1994*, Philosophy of Science Association, East Lansing, Michigan, vol. 2, 1995, pp. 156-164; HOWSON, C., «Accomodation, Prediction and Bayesian Confirmation Theory», en FINE, A. y LEPLIN, J. (ed.), *PSA 1988*, Philosophy of Science Association, East Lansing, Michigan, vol. 2, 1989, pp. 381-392; y HOWSON, C., «Fitting your Theory to the Facts: Probably not such a Bad Thing after All», en WADE SAVAGE, C. (ed.), *Scientific Theories*, Minnesota Studies in the Philosophy of Science, Minneapolis, 1990, pp. 224-244.

30. Cfr. BRUSH, S. G., «Dynamics of Theory Change: The Role of Predictions», en HULL, D., FORBES, M. y BURIAN, R. M. (eds.), *PSA 1994*, Philosophy of Science Association, East Lansing, vol. 2, 1995, pp. 133-145.

31. Sobre este tema, cfr. GONZÁLEZ, W. J., «Lakatos's Approach on Prediction and Novel Facts», pp. 505-508.

metodológica de la predicción sobre la explicación sigue siendo debatida, pero persiste la intuición en favor de dar más relevancia al acierto en el conocimiento del futuro —aquello que aún no es— que al proceso de acomodar nuestras explicaciones al presente y al pasado (aquello que es o ha sido, y posee una realidad que ya no va a cambiar en cuanto tal).

3. Tipos de explicaciones atendiendo a facetas metodológicas

Dentro de su estudio de las pautas de la explicación científica, E. Nagel mantuvo que había una conexión entre la pregunta *por qué*, como propia de la explicación científica, y la existencia de una *variedad* de explicaciones científicas, que —a su vez— llevaría a una de las diversas facetas metodológicas, en consonancia con los contextos temáticos. Así, a su juicio, las «explicaciones son las respuestas a la pregunta '¿por qué?'. Sin embargo, hace falta muy poca reflexión para mostrar que la palabra 'por qué' no es inequívoca (*unambiguous*) y que, al variar los contextos, caben diferentes tipos de contestaciones como respuestas relevantes ante ella».[32] Por tanto, el mismo tipo de pregunta —el interrogante propio de la «explicación científica»— ya hace presagiar la existencia de una diversidad de respuestas y, en consecuencia, un abanico de explicaciones posibles.

Este factor de variabilidad afecta a todas las Ciencias Empíricas, pero atañe singularmente a las Ciencias Humanas y Sociales. En ellas, incluso con mayor claridad que en las Ciencias de la Naturaleza, se han contrapuesto con gran frecuencia dos tipos de explicaciones: las causales y las teleológicas (habitualmente redenominadas en este caso como *intencionales*, para diferenciarlas de las meramente «funcionales», que también se darían en Ciencias de la Naturaleza, especialmente en Biología). Ahora bien, la existencia de diferentes respuestas ante la misma pregunta, a tenor del contexto, abriría la puerta a la compatibilidad entre «causas» y «fines» en el campo humano y social.[33] En tal caso, una acción social podría ser intencional bajo una descripción —y, por tanto, orientada a fines— y, al mismo tiempo, causal.

32. *The Structure of Science*, pp. 15-16.

33. Sobre causalidad y teleología en el campo humano y social, cfr. GONZÁLEZ, W. J., «From *Erklären-Verstehen* to *Prediction-Understanding*: The Methodological Framework in Economics», en SINTONEN, M., YLIKOSKI, P. y MILLER, K. (eds.), *Realism in Action: Essays in the Philosophy of Social Sciences*, Kluwer, Dordrecht, 2002, pp. 38-40.

3.1. Cuatro tipos de explicación científica

Nagel propuso una distinción entre cuatro tipos de explicación científica —antes mencionados—, haciendo una síntesis representativa de las aportaciones realizadas hasta ese momento. En cada uno de los tipos se asume que hay *un hecho* que se quiere explicar o *explicandum*: algo que acaece siempre; o sólo con una determinada cadencia; o que depende de un fin al que se orienta; o tiene un específico desarrollo en el tiempo. Para responder a la correspondiente pregunta «por qué», hace falta un *explicans*, que se concibe entonces como la argumentación a partir de la cual se da razón del *explicandum*. Son, por tanto, en su enfoque cuatro pautas lógicas (*logical patterns*) que concibe como diferentes.[34]

Destaca *el modelo deductivo*, al que dedica un amplia atención.[35] Lo concibe como una argumentación donde el hecho a explicar (*explicandum*) es una consecuencia lógica necesaria de las premisas explicativas. Advierte que no toda explicación resulta aceptable a tenor de la mera posesión de una estructura deductiva (p. ej., no es satisfactorio como explicación de un satélite en Júpiter que se mencione el hecho de tener ocho lunas, aunque lo primero se pueda seguir lógicamente de lo segundo). A su juicio, hace falta el cumplimento de condiciones de diversa índole: i) *lógicas*, que especifican los requisitos formales que han de cumplir las premisas explicativas; ii) *epistémicas*, que indican las relaciones cognitivas que deberían poseer las premisas, y iii) *sustantivas*, que señalan un tipo de contenido (p. ej., semántico) que deberían tener las premisas.[36] De ellas son particularmente conocidas las condiciones lógicas: cuando se trata de explicar un suceso particular, hace falta una *ley universal* en las premisas (para deducir el *explicandum*) y una serie de *condiciones iniciales* que especifiquen el caso (p. ej., la imposibilidad de arrancar un vehículo por las bajas temperaturas en una noche concreta).[37]

Frente al modelo deductivo contrapone el tipo de *las explicaciones probabilísticas*. Lo concibe como una variedad explicativa ajena a la forma deductiva, de modo que las premisas explicativas no implican formalmente aquello que desean explicar, es decir, la estructura argumental no cuenta con premisas que aseguren la verdad del *explicandum*, pero lo pueden hacer «probable». Considera,

34. Cfr. Nagel, E., *The Structure of Science*, p. 21.
35. Cfr. *The Structure of Science*, pp. 29-46.
36. Cfr. Nagel, E., *Ibidem*, p. 30.
37. Cfr. *The Structure of Science*, p. 32.

además, que es habitual que las premisas contengan una suposición estadística acerca de algunas clases de elementos y que el *explicandum* sea un enunciado singular acerca de un determinado miembro de la clase estudiada. Es algo que considera ejemplificable en las Ciencias Humanas y Sociales (p. ej., la explicación del asesinato de Julio César a partir de una alta probabilidad, entendida en términos de frecuencia relativa, asociada a miembros de su posición social).[38] Según lo plantea Nagel, se trata de un tipo de explicación distinto al deductivo, no meramente un paso temporal para llegar al «ideal» del modelo deductivo.[39]

Distinto de los otros dos tipos es la explicación científica *funcional* o *teleológica*. Mira hacia un estado futuro o evento en términos del cual resulta inteligible la existencia de una cosa o la realización de un acto. Sus expresiones características son «con el fin de», «con el propósito de» y muchas otras de esta índole. Suele emplearse en Biología y en Ciencias Humanas y Sociales, aunque no sólo en ellas. Puede, a su vez, diversificarse en dos posibilidades: *a*) la explicación funcional de un acto, estado o cosa particular que acaece en un determinado momento (p. ej., una actuación con un fin determinado, como la búsqueda de la anulación del matrimonio de Enrique VIII); y *b*) la explicación funcional de un rasgo presente en sistemas de una cierta clase, sea cual fuere el momento (p. ej., la existencia de un órgano en todos los sistemas de unas ciertas características, al margen del momento específico, como la presencia de los pulmones en el cuerpo humano). Nagel considera, además, que es un error tanto asociar las explicaciones teleológicas sólo a agentes conscientes como también el vincularlas a un presunto carácter causal del futuro sobre el presente.[40]

Respecto de las *explicaciones genéticas*, Nagel remite a la dimensión histórica: las investigaciones históricas tratan con fre-

38. Cfr. NAGEL, E., *The Structure of Science*, pp. 22-23.

39. Georg H. von Wright marca mucho las diferencias: «The two models are much more different than is often thought. It is a primary function of the deductive-nomological model to explain why certain things happened. Therefore, it also tells us, secondarily, why these things were to be expected. They could have been expected, *since* they had to happen. With the inductive-probabilistic model the roles are reversed. It explains in the first place why things which happened were to be expected (or not to be expected). Only in a secondary sense does it explain why things happened, *viz.* 'because' they were highly probable. It seems to me better, however, not to say that the inductive-probabilistic model explains what happens, but to say only that it justifies certain expectations and predictions», WRIGHT, G. H. VON, *Explanation and Understanding*, Cornell University Press, Ithaca, 1971, p. 14.

40. Cfr. *The Structure of Science*, pp. 23-25.

cuencia de explicar la evolución de las características de alguna realidad, sea en el ámbito de los seres animados o bien en los inanimados, tanto en individuos como en grupos. Las explicaciones genéticas se ocupan de reconstruir la secuencia de los sucesos principales por los cuales un sistema originario ha dado lugar a otro distinto a través del tiempo (p. ej., por qué la lengua inglesa actual tiene tantas palabras de origen latino, especialmente cuando se trata de términos intelectuales). Las premisas explicativas remiten a enunciados singulares acerca de los acontecimientos del pasado que pertenecen al devenir del sistema investigado. En tal caso, habría que tener en cuenta varios aspectos: 1) en las premisas no se menciona todo suceso del pasado; 2) los sucesos elegidos suelen serlo por su importancia causal para el sistema, y 3) suele haber supuestos generales acerca de las dependencias causales entre diferentes grupos de eventos. Estos supuestos pueden ser o bien precisos (p. ej., a tenor de leyes de desarrollo apoyadas empíricamente, como la evolución de características biológicas de un miembro individual de una especie), o bien genéricos (p. ej., cuando se investiga la trayectoria de alguna institución en una cultura particular).[41]

3.2. La variedad de tipos explicativos en W. C. Salmon

Comparando el esquema cuatripartito de los tipos de explicación científica de Nagel y las propuestas realizadas por Salmon, parece claro que los dos primeros tipos —el «modelo deductivo» y las «explicaciones probabilísticas»— han sido objeto de su análisis durante décadas, llegando a desarrollos mucho más sofisticados que su ilustre antecesor. También las «explicaciones funcionales» hacen acto de presencia en sus escritos. Lo hace con cautela: acepta la existencia de explicaciones funcionales en Biología, Antropología y Sociología, pero piensa que pueden ser interpretadas como un subconjunto de las explicaciones causales.[42] Ve con recelo el entenderlas en términos teleológicos (esto es, relacionadas con motivos, propósitos o fines), pues afirma que «históricamente tanto la Física como la Biología han hecho progresos significativos al eliminar las explicaciones teleológicas».[43]

41. Cfr. Nagel, E., *The Structure of Science*, pp. 25-26.
42. Cfr. Salmon, W. C., *Causality and Explanation*, p. 7.
43. *Causality and Explanation*, p. 7.

No obstante, deja una puerta abierta a esa faceta al tratar de la intencionalidad de la acción humana: «aunque he trazado antes una clara distinción entre las explicaciones de significados (*meanings*) y las explicaciones de fenómenos naturales, a veces se puede decir que una comprensión adecuada de la conducta humana debe incluir las interpretaciones de los significados. La razón es que muchos actos se llevan a cabo debido a su significado. Los ritos religiosos, por ejemplo, serían a menudo ininteligibles desprovistos de su significado simbólico».[44] En cambio, las «explicaciones genéticas» difícilmente encuentran acogida expresa en sus trabajos, salvo en cuanto que pueden retrotraerse a explicaciones causales o en la medida en que son probabilísticas.[45]

Cabe reflejar con claridad su posición sobre los tipos explicativos gracias a su honradez intelectual, que ha caracterizado toda su trayectoria académica. Así, en la parte conclusiva de su estudio acerca de cuatro décadas de la explicación científica, W. C. Salmon indica expresamente que en el libro «han quedado virtualmente sin tocar algunas amplias áreas. He dejado de lado de modo deliberado temas tales como la explicación en Biología, la explicación en Psicología, la explicación en Historia y la explicación de la acción humana, en gran medida porque no poseo la información suficiente acerca de su contenido fundamental. Algunas cuestiones relacionadas, especialmente la causalidad, han sido tratadas aquí sólo de pasada, porque la literatura es muy amplia y también porque este ensayo es ya demasiado largo».[46] Pero la causalidad y su nexo con la explicación científica aparecen tratadas *in extenso* tanto en su libro *Causality and Explanation* como en los trabajos del presente volumen, sobre todo el segundo.[47] En cambio, el dominio de la explicación en las Ciencias Humanas y Sociales pocas veces aparece en sus escritos, siendo Merrilee H. Salmon quien ha proyectado sus ideas en ese ámbito.[48]

Así pues, como se puede apreciar, W. C. Salmon se ha centrado en las Ciencias de la Naturaleza, en especial en la Física. Sólo tangencialmente alude a las Ciencias Humanas y Sociales, pero apenas hay desarrollo alguno de las explicaciones en el ám-

44. SALMON, W. C., *Ibidem*, pp. 7-8.

45. Para Nagel, «it is (...) a reasonable conclusion that genetic explanations are by and large probabilistic», *The Structure of Science*, p. 26.

46. *Four Decades of Scientific Explanation*, p. 180.

47. Cfr. SALMON, W. C., «Estructura de la explicación causal», pp. 141-159.

48. Cfr. SALMON, M. H., «La explicación causal en Ciencias Sociales», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 161-180.

bito que Herbert A. Simon ha llamado «Ciencias de lo Artificial».[49] El campo de las Ciencias de Diseño requiere también atención y puede dar lugar —a mi juicio— a nuevos tipos de explicaciones científicas, pues el tipo de pregunta «por qué» remite a un contexto diferente. Parece *prima facie* que habría una diferencia con el tipo de explicaciones que Salmon señala para las Ciencias Empíricas: propiamente las explicaciones no serían de «hechos» sino de *acciones humanas* encaminadas a lograr objetivos (*aims*) deliberadamente pensados.

Básicamente, los tipos de explicaciones científicas admitidos por Salmon son el resultado de su interacción con las diversas etapas que estudia en su célebre libro.[50] Así, reconstruir su postura supone partir de un primer enfoque de la explicación científica, previo a la consolidación de una postura dominante, que llama «Concepción heredada» (*received view*) acerca de la explicación científica. Corresponde al «modelo nomológico-deductivo», donde hay leyes —*nomos*— y deducciones en lugar de inducciones. Más tarde, se propuso esa posición dominante, donde la explicación científica debía satisfacer *dos características* decisivas: primera, debe ser o un argumento deductivo válido o un argumento inductivo correcto; y, segunda, debe contener una ley universal.

Mediante la *Concepción heredada* aparecieron cuatro modelos: 1) las explicaciones *nomológico-deductivas de regularidades universales* (p. ej., la ley de conservación del momento, que se explica a través de una inferencia deductiva a partir de las tres leyes de la Mecánica newtoniana); 2) las explicaciones *nomológico-deductivas de hechos particulares* (p. ej., la vuelta del cometa Halley en 1759, explicada mediante deducción desde las leyes de gravitación y del movimiento de Newton junto con las condiciones iniciales de la configuración de los cuerpos en el Sistema Solar); 3) las explicaciones *estadístico-deductivas de regularidades estadísticas* (p. ej., que haya menos de un tanto por ciento de posibilidades de sacar el seis doble en un número concreto de tiradas de un par de dados normales, que se deduce de leyes probabilísticas, características simétricas de un dado legítimo y leyes físicas de los sólidos simétricos), y 4) las *explicaciones estadístico-inductivas de hechos particulares*

49. Cfr. SIMON, H., *The Sciences of the Artificial*, 3.ª ed., M.I.T. Press, Cambridge, 1996 (1.ª ed., 1969; 2.ª ed., 1981).

50. Las posiciones más relevantes son analizadas en SALMON, W. C., «Explicación causal frente a no causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 97-115. En lo que sigue se atiende expresamente a esa presentación.

(p. ej., que una persona concreta no adquiera una enfermedad, gracias a una vacuna, cuando muchas personas del entorno se hubieran contagiado).[51]

Esos diferentes modelos de explicación científica suscitaron una serie de contraejemplos, que acabaron dañando su credibilidad. Esto le llevó a Salmon, habitualmente crítico con la concepción de Hempel,[52] a proponer en los años setenta una alternativa para la explicación estadística.[53] El requisito central era la *relevancia estadística* en vez de la *alta probabilidad*. Se comienza, por ejemplo, con un caso concreto que pertenece a un grupo general de referencia, como los infectados por una bacteria (p. ej., el estreptococo). Puede haber una persona concreta que se recobrara rápidamente, y preguntamos *por qué*. A tenor del modelo de relevancia estadística, se hace una división relevante del grupo de referencia en dos subgrupos: por un lado, quienes fueron tratados con penicilina y, por otro, aquellos que no recibieron el tratamiento. La probabilidad de una pronta recuperación es mucho mayor en el primer subgrupo (el tratado con penicilina) que en el conjunto del grupo de referencia o en el segundo subgrupo (aquellos que no recibieron el tratamiento). La *relevancia estadística* estaría, por tanto, en que el tratamiento médico marca la diferencia en la probabilidad del resultado.

Posteriormente, Salmon vio que este planteamiento suponía mejoras respecto del modelo estadístico inductivo, pero resultaba una posición inadecuada en cuanto que se encamina hacia la relevancia estadística en lugar de resaltar la relevancia causal. Esto podría traer como resultado el llegar a confundir «causas» con «correlaciones». Considera entonces otras posibilidades: *a*) el *modelo de unificación*, que atañe a generalizaciones y supone que se puede unificar un gran número de regularidades específicas en una teoría mediante un número reducido de supuestos o postulados;[54] *b*) el enfoque de la *tradición mecánico-causal*, donde la explicación causal requiere descubrir los mecanismos que originan los hechos que

51. Salmon apunta que Hempel y Oppenheim vieron en 1948 las explicaciones nomológico-deductivas, pero que sólo desarrollaron en rigor la segunda (hechos particulares). Después, en 1962, Hempel ofreció el esquema estadístico-inductivo y en 1965 trató los enfoques estadístico-deductivo y estadístico-inductivo. Cfr. SALMON, W. C., «Scientific Explanation», p. 17.

52. Cfr. FETZER, J. H. (ed.), *Science, Explanation, and Rationality: The Philosophy of Carl G. Hempel*, Oxford University Press, N. York, 2000, p. xxv.

53. Cfr. SALMON, W. C. (ed.), *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh, 1971.

54. Cfr. FRIEDMAN, M., «Explanation and Scientific Understanding», *Journal of Philosophy*, v. 71, (1974), pp. 5-19.

buscamos comprender,[55] y *c*) la *concepción pragmática*, que insiste en el énfasis que pone el hablante como crucial para determinar qué cuestión se está planteando.[56] Entiende que cada una de ellas ha de encarar problemas de fondo,[57] por lo que opta por resaltar la relación entre explicación y causalidad.

4. La preferencia por la explicación causal

Una vez analizadas otras opciones, W. C. Salmon pone de relieve una inequívoca preferencia por las explicaciones científicas causales. Esto ha motivado, a su vez, una atención especial a la noción de «causa» y a la crítica que realizó David Hume. Su posición es clara: «sin insistir en que todas las explicaciones científicas son causales, podemos no obstante mantener que el conocimiento de las relaciones causales nos permite explicar un amplio número de fenómenos naturales, y que esas explicaciones proporcionan una comprensión del mundo y de lo que acaece en él».[58] Esa afirmación, escrita justo después de terminar la quinta década de la explicación científica (1988-1997), sirve de marco para entender sus reflexiones de este último periodo.

Resulta sintomático que su primera contribución a este volumen termine con una apelación: tanto la explicación como la causalidad necesitan una base enteramente objetiva. Por un lado, esto supone que, sin negar aspectos pragmáticos en la explicación científica, la clave no está ahí. La base objetiva conlleva distinguir entre explicaciones de hechos particulares y explicaciones de regularidades generales. Más aún, le parece razonable creer que, aun cuando hay dificultades, podemos explicar las regularidades generales deduciéndolas a partir de leyes generales de nivel más alto (p. ej., las tres leyes del movimiento de Newton explican por qué se ha de conser-

55. La consideración de la causalidad tiene un papel relevante en la crítica a Hempel y Oppenheim realizada en SCRIVEN, M., «Explanations, Predictions and Laws», en FEIGL, H. y MAXWELL, G. (eds.), *Scientific Explanation, Space and Time*, University of Minnesota Press, Minneapolis, 1962, pp. 170-230. Sobre el tema de la causalidad versa también SCRIVEN, M., «Causation as Explanation», *Nous*, v. 9, (1975), pp. 3-16.

56. Cfr. FRAASSEN, B. C. VAN, *The Scientific Image*, Clarendon Press, Oxford, 1980. Sobre el papel de la pragmática en la explicación científica hay también reflexiones en RAILTON, P., «Probability, Explanation, and Information», *Synthese*, v. 48, (1981), pp. 233-256.

57. Además de las observaciones críticas aportadas en su primer trabajo en este volumen —«Explicación causal frente a no causal»—, cfr. SALMON, W. C., «Scientific Explanation», pp. 33-37.

58. *Causality and Explanation*, p. 8.

var el momento lineal). Y, por otro lado, afirma la existencia del nexo causal y la vinculación entre explicaciones de hechos particulares y relaciones causales (p. ej., explicar por qué se estrella un avión estriba en dar con la causa o las causas del accidente).[59]

Según Salmon, hace falta proponer enfoques radicalmente diferentes de los conceptos de «explicación» y de «causalidad», en especial, de *causalidad*. Por eso dedica su segundo trabajo de este volumen a proponer «una teoría satisfactoria de la causalidad». A partir de ahí espera ofrecer una base sólida para una *teoría causal de la explicación científica*. Es una posición que se contrapone a los planteamientos no causales de la explicación científica, tales como la «Concepción heredada», cuya versión más pulida se encuentra en C. G. Hempel, y que fue asumida por muchos autores. Es también una concepción diferente del modelo de la relevancia estadística, que el propio Salmon propuso en los años setenta, y distinta también de otras posturas, como la propuesta de unificación de Michael Friedman o la teoría pragmática de Bas C. van Fraassen.

Propone, en suma, Salmon un nuevo análisis de la causalidad que sirva, a su vez, para una nueva visión de la explicación científica. De hecho, la articulación de ese segundo trabajo del volumen va en esa dirección: 1) un nuevo punto de partida, distinto del habitual, que le lleva a centrarse en la noción de «proceso»; 2) la separación entre procesos causales y pseudo-procesos; 3) los rasgos de la transmisión causal; 4) las interacciones causales (donde su cometido es caracterizar las interacciones causales en términos de los conceptos no causales de «proceso», «intersección» y «cambio»); 5) las relaciones entre las causas y los efectos (donde el modelo de causa-efecto consiste en una *interacción causal* entre dos procesos, un *proceso causal* que surge a partir de la interacción, y una *interacción* posterior de ese proceso con otro proceso); 6) la estructura causal completa (estructura que, para cualquier región espacio-temporal, se da mediante la red completa de procesos causales y de interacciones causales contenidas en esa región elegida), y 7) la objetividad del enfoque que propone.[60]

Hacia mediados de la cuarta década, en 1984, a través del libro *Scientific Explanation and the Causal Structure of the World* y de un conocido trabajo,[61] Salmon había dado a conocer su distinción de

59. Cfr. SALMON, W. C., «Explicación causal frente a no causal», p. 112.
60. Cfr. SALMON, W. C., «Estructura de la explicación causal», pp. 141-159.
61. SALMON, W. C., «Scientific Explanation: Three Basic Conceptions», *Proceedings of the Philosophy of Science Association*, v. 2, (1984), pp. 293-305.

tres visiones generales de la explicación científica: la *epistémica*, la *modal* y la *óntica*. La postura epistémica corresponde a la «Concepción heredada» de Hempel. El planteamiento modal ofrece una visión distinta: explicar significa mostrar que lo que *de hecho* sucedió *tenía que* suceder a la vista de los hechos explicativos. Y la opción óntica, a diferencia de la posición epistémica, considera que las explicaciones son completamente objetivas (esto es, no son relativas a nuestra situación de conocimiento). Desde luego, esta posición —apoyada por nuestro autor— supone que depende de nuestro estado de conocimiento lo que nosotros *consideramos que son* explicaciones correctas.

Continuando con esa opción, en el segundo trabajo del presente libro Salmon profundiza en favor de una interpretación óntica de la causalidad. Así, al analizar la *estructura causal completa* en términos de interacciones (causales), transmisiones (causales) y procesos causales, intenta extender el fundamento (*groundwork*) objetivo de la causalidad. Al mismo tiempo, sostiene que las afirmaciones de causa-efecto dependen habitualmente (si no siempre) de consideraciones contextuales. Por eso, entiende que es un error mantener que hemos encontrado *la única* explicación correcta de un hecho dado, y considera que es asimismo un error defender que hemos encontrado *la única* causa de un efecto dado. Piensa, por último, que identificar una causa es, en muchas situaciones, equivalente a dar una explicación. Su cometido principal ha sido proporcionar un enfoque de la causalidad que haga justicia a los aspectos objetivos y de dependencia del contexto de la causalidad y de la explicación (y, paralelamente, haga justicia también a la crítica clásica de Hume a la causalidad).

Importa resaltar que, para Salmon, «sería un serio error suponer que cada fenómeno tiene una única explicación. Creo que es una equivocación preguntar por *la* explicación de algo que ocurre».[62] Esto abre la puerta a la complementariedad entre explicaciones. Así, en 1992 advertía que durante más de treinta años (a contar desde 1959) ha habido «una fuerte oposición entre los partidarios de la Concepción heredada y los proponentes de la explicación causal. Cada uno de los enfoques principales ha evolucionado de manera considerable durante este periodo; más aún, se han desarrollado hasta el punto de poder coexistir pacíficamente como dos aspectos distintos de la explicación científica. La comprensión científica es, después de todo, un asunto complicado; no

62. SALMON, W. C., «Scientific Explanation», p. 38.

deberíamos sorprendernos de aprender que tiene muchos aspectos diferentes».[63]

Visto en retrospectiva, la trayectoria de W. C. Salmon acerca de las explicaciones científicas combina las dos vertientes señaladas al principio: por un lado, intenta aclarar la caracterización de las variedades legítimas de «explicación científica», inclinándose por la respuesta dada a ciertas formas de usar la pregunta *por qué* (sin descartar respuestas ante preguntas sobre *cómo* algo ha sido posible o se ha formado); y, por otro lado, busca explicitar los tipos de explicaciones científicas que plasman aspectos presentes en el proceso metodológico de la Ciencia, decantándose por las *explicaciones causales*, pero admitiendo su complementariedad con otras posibilidades.

5. Origen y estructura del presente volumen

Tiene su origen el presente volumen en las *Jornadas sobre Variedades de explicación científica. El enfoque de W. Salmon*, que se celebraron los días 8 y 9 de marzo de 2001 en la Universidad de A Coruña, Campus de Ferrol. Eran las *VI Jornadas sobre Filosofía y Metodología actual de la Ciencia*, organizadas por la Facultad de Humanidades en colaboración con la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España.[64] En este caso, el Congreso estaba centrado en un problema básico: los planteamientos actuales acerca de la explicación científica, viendo las variedades existentes. Se buscaba sobre todo la consideración de las posiciones correspondientes a la larga «quinta década» de la evolución temática de la explicación científica, una vez que Wesley Salmon había analizado las anteriores cuatro etapas (1948-1987). A este respecto, había

63. «Scientific Explanation», pp. 38-39. En el transcurso de una pregunta, realizada el 9.3.2001, sobre la compatibilidad entre explicaciones causales y explicaciones intencionales, W. C. Salmon aceptó la posibilidad de combinar ambos aspectos.

64. Los principales trabajos de las anteriores Jornadas están publicados dentro de una colección general que lleva por título *Gallaecia. Estudios de Filosofía y Metodología de la Ciencia* y lo componen los siguientes volúmenes: GONZÁLEZ, W. J. (ed.), *Progreso científico e innovación tecnológica*, número monográfico de *Arbor*, v. 157, n. 620, (1997); GONZÁLEZ, W. J. (ed.), *El Pensamiento de L. Laudan. Relaciones entre Historia de la Ciencia y Filosofía de la Ciencia*, Publicaciones Universidad de A Coruña, A Coruña, 1998; GONZÁLEZ, W. J. (ed.), *Ciencia y valores éticos*, número monográfico de *Arbor*, v. 162, n. 638, (1999); GONZÁLEZ, W. J. (ed.), *Problemas filosóficos y metodológicos de la Economía en la Sociedad tecnológica actual*, número monográfico de *Argumentos de Razón técnica*, v. 3, (2000); y GONZÁLEZ, W. J. (ed.), *La Filosofía de Imre Lakatos: Evaluación de sus propuestas*, UNED, Madrid, 2001.

sido elegido como invitado principal por su reconocido prestigio internacional como autoridad en el estudio histórico y sistemático de los tipos de explicaciones científicas.

Para articular temáticamente los contenidos, en el presente volumen se han resaltado cuatro líneas principales, que dan lugar a las partes señaladas en el índice general: «Evolución del problema de la explicación científica»; «Propuestas para las explicaciones científicas»; «Las explicaciones causales: estructura y aplicación»; y «Explicaciones teleológico-causales y explicaciones programáticas». Cada una de ellas está representada por dos trabajos. Además de contemplar el *nivel de generalidad* de cada estudio, al diseñar esa estructura del libro se ha tenido en cuenta la *precedencia en el tiempo* de los diversos planteamientos. Prevalece así el vector temático —su contribución a la especificación de las variedades de explicación científica— pero sin desatender el plano histórico.

Subyace a la estructura del libro el problema inicial, esto es, la existencia de los dos planos señalados al comienzo. Por un lado, se considera la caracterización de los *rasgos propios* de la «explicación científica»; cuestión que, en buena medida, depende del énfasis que se ponga en uno u otro de los componentes de la Ciencia (el factor de conocimiento bien fundado, la dimensión de actividad humana, ...) y del enfoque general adoptado al respecto (argumento, enunciado, pregunta, ...). Y, por otro lado, se atiende al contenido a través de la *faceta metodológica* que es resaltada en la explicación propuesta (es decir, su nexo con causas, fines, ...). Esto modula la orientación de la explicación científica (causal, teleológica, ...), lo que, a su vez, repercute en diversos planos (en las Ciencias Empíricas en su conjunto; en las Ciencias de la Naturaleza o, en su caso, en las Ciencias Humanas y Sociales; en una Ciencia concreta; ...). Así, aun cuando caben siempre matices que son relevantes, se puede afirmar que hay mayor atención a la primera vertiente en los apartados 2 y 3 del volumen, mientras que se centran claramente en la segunda vertiente los epígrafes 4 y 5 del libro.

En efecto, al tratar de la *Evolución del problema de la explicación científica*, se atiende tanto al pasado como al presente: se analiza la adaptación histórica de la explicación científica al problema de la complejidad (Javier Ordóñez) y se constata la existencia de diversos modos de entender la explicación científica en cuanto tal (José A. Díez Calzada). A continuación, al estudiar las *Propuestas para las explicaciones científicas*, se examina la trayectoria seguida en la segunda mitad del siglo XX por diferentes tipos de explicaciones científicas, al hilo de las posiciones sobre la explicación causal

y teniendo como eje conductor la explicación de hechos (Wesley C. Salmon), y se expone un planteamiento articulado en torno a la Ciencia como actividad humana y sustentado en la relevancia de valores científicos (Javier Echeverría).

Más tarde la atención se concentra en tipos específicos de explicación, con especial insistencia en la orientación causal. Así, al considerar *Las explicaciones causales: estructura y aplicación*, el énfasis se pone en una modulación concreta de las explicaciones científicas: la causal, que afecta, en principio, al conjunto de la Ciencias Empíricas, apreciándose su presencia tanto al tratar de las Ciencias de la Naturaleza (Wesley C. Salmon) como al abordar las Ciencias Humanas y Sociales (Merrilee H. Salmon). Posteriormente, al contemplar el campo de *Explicaciones teleológico-causales y explicaciones programáticas*, la mirada se dirige hacia las concepciones específicas que atañen especialmente a las Ciencias Humanas y Sociales, viendo cómo articular la aceptación de la finalidad y la causalidad (Rafaela García Elskamp) y cómo relacionar las explicaciones funcionales con las programáticas (David Pineda).

Cuando se alcanza el término del proceso que se inició con la preparación de las Jornadas —que culmina con la publicación del presente volumen—, resulta natural la gratitud hacia quienes lo han hecho posible o han colaborado en algún momento. El agradecimiento se dirige entonces al Ministerio de Ciencia y Tecnología, por el respaldo dado por la Dirección General de Investigación a esta iniciativa. Reconocimiento asimismo al Concello de Ferrol, tanto al Alcalde como al Concelleiro de Cultura y Campus, quienes han resaltado la relevancia de este vínculo entre Universidad y Sociedad; un entrelazamiento que, en este caso, no constituye un *desideratum* buscado, sino una realidad. Gratitud también hacia el Banco Pastor y Caixa Galicia, que contribuyen a las actividades promovidas por el Vicerrectorado de Relaciones Institucionales y Postgrado, que —como años anteriores— ha propiciado la realización de estas Jornadas.

Por lo que hace al entorno inmediato, mi agradecimiento a la Facultad de Humanidades en estas *VI Jornadas de Filosofía y Metodología actual de la Ciencia* se personaliza en el Decano, Víctor Alonso Troncoso, atento siempre a promover e impulsar actividades de alto nivel intelectual en la Facultad. También deseo dar las gracias a quienes han ayudado en el ámbito material y organizativo: el grupo de alumnos y el personal administrativo de la Facultad. Asimismo quiero resaltar el papel de los medios de comunicación del entorno, que han puesto de relieve el interés que tienen estas Jornadas.

Obviamente, mi gratitud cobra particular relieve en el caso del profesor Wesley C. Salmon, por haber aceptado la invitación a participar en las Jornadas y por haber seguido las sugerencias que le hice desde 1999 respecto de los temas de las ponencias, que ha permitido tener este legado acerca de la explicación científica. También a Merrilee H. Salmon, Catedrática de la Universidad de Pittsburgh, que ha sido editora de *Philosophy of Science*, le expreso mis sinceras gracias por su participación en las Jornadas y en este volumen, y por haberme facilitado datos de interés, especialmente tras el fallecimiento de Wesley Salmon. Asimismo reitero mi agradecimiento a los demás ponentes del Congreso —Javier Echeverría, Javier Ordóñez y José Antonio Díez Calzada— y a quienes presentaron Comunicaciones. Doy las gracias también a José Francisco Martínez Solano, por su colaboración en las tareas de edición.

6. Trayectoria académica

Pocas semanas después de haber participado en las Jornadas celebradas en su honor en la Universidad de A Coruña, el domingo 22 de abril de 2001, tenía lugar lo inesperado: en Madison County (Ohio) fallecía en accidente de tráfico Wesley Charles Salmon. Iba acompañado por Merrilee H. Salmon, que no sufrió daños físicos. Habían salido de Indiana hacia Pensilvania. Se desplazaban en automóvil en dirección a Pittsburgh, cuando su coche fue alcanzado por otro vehículo, falleciendo de forma instantánea.

Wesley C. Salmon era Catedrático emérito de Filosofía de la Universidad de Pittsburgh y también *Professor Emeritus* del Departamento de Historia y Filosofía de la Ciencia de la misma Universidad. Durante 15 años (1983-1999) ocupó la Cátedra que había tenido Carl G. Hempel, otra de las grandes figuras de la Filosofía de la Ciencia. Gozaba de reconocimiento mundial y había sido el Presidente de la Sociedad Internacional de Lógica, Metodología y Filosofía de la Ciencia (1996-99) y fue también Presidente de la Unión Internacional de Historia y Filosofía de la Ciencia (1998-99).

Salmon había nacido el 9 de agosto de 1925 en Detroit (Michigan). Se había formado en la Universidad de Chicago, donde obtuvo la Licenciatura en 1947, haciendo ese año y en esa misma Universidad un Máster en Filosofía. A continuación se trasladó a la Universidad de California en Los Ángeles (UCLA), donde realizó los estudios de Doctorado, presentando la Tesis Doctoral en 1950. Su Director —al igual que en el caso de H. Putnam— fue Hans Reichenbach,

siendo por tanto discípulo directo suyo.[65] Tras impartir docencia en la UCLA, fue profesor de las Universidades de Washington (State College), Northwestern, Brown, Indiana, Arizona y Pittsburgh. También fue profesor visitante en las Universidades de Bristol, Chicago, Minnesota, Melbourne, Bolonia, Constanza y Kioto.

Su prestigio le llevó a ocupar importantes puestos: había sido Presidente de la *Philosophy of Science Association* (1971-72); Presidente de la *American Philosophical Association*, Pacific Division (1977-78); y Presidente de la *International Union of History and Philosophy of Science* (1998-99). También había sido miembro de importantes comités editoriales: *American Philosophical Quarterly* (1964-67); *Erkenntnis* (1975-91); *Journal of Philosophical Logic* (1970-84); *Pacific Philosophical Quarterly* (1979-2001); *Philosophy of Science* (1962-75); *Synthese* (1966-88); *Synthese Library* (1966-93); y *Pittsburgh Series in Philosophy and History of Science*. Escribió libros y artículos realmente influyentes, tanto para el tema de la explicación científica como para el problema de la causalidad; títulos cuyas referencias completas se detallan en el siguiente epígrafe, al exponer su bibliografía.

Del paso de Wesley Salmon por nuestro país —en especial, de su estancia en la Facultad de Humanidades de Ferrol—, además del grato recuerdo de su calidad humana y de su participación activa en las Jornadas sobre el tema en que era un reconocido especialista —la explicación científica—, quedará también el testimonio escrito de las dos ponencias que presentó. En efecto, esos dos trabajos son los que forman parte de este volumen en su honor y cabe verlos como el colofón inesperado de una larga y fructífera trayectoria profesional.

Personalmente, sólo tengo palabras de gratitud hacia Wesley C. Salmon, tanto por las atenciones recibidas en mis estancias en el *Center for Philosophy of Science* de la Universidad de Pittsburgh (1993-94, 1996 y 1999) como por la ayuda que me ha prestado cada vez que se lo he pedido. A este respecto, he de resaltar las conversaciones sobre el problema de la predicción científica y acerca de diversos aspectos del Empirismo lógico, que fueron de indudable utilidad para mi análisis de la predicción en Hans Reichenbach.[66]

65. El reconocimiento de H. Reichenbach como su Maestro se hace patente en el libro homenaje SALMON, W. C. (ed.), *Hans Reichenbach: Logical Empiricist*, Reidel, Dordrecht, 1979.

66. Cfr. GONZÁLEZ, W. J., «Reichenbach's Concept of Prediction», *International Studies in the Philosophy of Science*, v. 9, n. 1, (1995), pp. 35-56.

7. Bibliografía de Wesley C. Salmon

Gran parte de la bibliografía de W. C. Salmon está relacionada de manera directa o indirecta con el problema de la explicación científica. Por eso, se ha optado por ofrecer un amplio elenco de sus trabajos académicos, de modo que, primero, se ofrecen las referencias de sus libros y, después, se dan los datos de sus artículos y capítulos de libros, al tiempo que se dejan al margen otras publicaciones (réplicas, reseñas, colaboraciones en enciclopedias filosóficas, presentaciones de libros, ...). Como complemento, para cubrir el ámbito aquí abordado, se hace más tarde una enumeración de los volúmenes colectivos que analizan expresamente su concepción filosófica.

7.1. Libros

Se enumeran, a continuación, todos los libros que escribió o editó Wesley C. Salmon, tanto los que preparó de modo individual como los elaborados en colaboración con otros autores. En esta relación bibliográfica se da prioridad al orden cronológico, de modo que los trabajos compartidos aparecen a tenor de su fecha de publicación en lugar de hacerlo en función del apellido.

SALMON, W. C., *Logic*, Prentice Hall, Englewood Cliffs, NJ, 1963 (2.ª ed. 1973; 3.ª ed. 1984).[67]

SALMON, W. C., *The Foundations of Scientific Inference*, University of Pittsburgh Press, Pittsburgh, 1967.[68]

SALMON, W. C. (ed.), *Zeno's Paradoxes*, Bobbs-Merrill, Indianápolis, 1970 (reimpreso con nuevo prefacio en Hackett, Indianápolis, 2001).

SALMON, W. C. (ed.), *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh, 1971.[69]

SALMON, W. C., *Space, Time and Motion*, Dickenson, Encino, 1975; 2.ª ed. revisada, 1981.

67. Ha sido un libro traducido a diversas lenguas: castellano, japonés, italiano, portugués y alemán.

68. Se publicó inicialmente como un capítulo de libro: SALMON, W. C., «The Foundations of Scientific Inference», en COLODNY, R. G. (ed.), *Mind and Cosmos*, University of Pittsburgh Press, Pittsburgh, 1966, pp. 135-275. Fue reimpreso como libro en 1967 con un anexo. Más tarde, partes de este libro (pp. 5-27, 40-43, 48-56 y 132-136) fueron publicadas en una compilación: SALMON, W. C., «The Problem of Induction», en PERRY, J. y BRATMAN, M. (eds.), *Introduction to Philosophy*, Oxford University Press, N. York, 1986, pp. 265-286.

69. Cuenta con contribuciones de Richard C. Jeffrey y James G. Greeno.

SALMON, W. C. (ed.), *Hans Reichenbach: Logical Empiricist*, Reidel, Dordrecht, 1979.

SALMON, W. C., *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, NJ, 1984.

GRÜNBAUM, A. y SALMON, W. C. (eds.), *The Limitations of Deductivism*, University of California Press, Berkeley, 1988.

KITCHER, PH. y SALMON, W. C. (eds.), *Scientific Explanation*, University of Minnesota Press, Minneapolis, 1989.

SALMON, W. C., *Four Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis, 1990.[70]

SALMON, M. ET AL., *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs, NJ, 1992.[71]

SALMON, W. C. y WOLTERS, G. (eds.), *Logic, Language, and the Structure of Scientific Theories*, University of Pittsburgh Press, Pittsburgh, 1994.

SALMON, W. C., *Causality and Explanation*, Oxford University Press, Nueva York, 1998.

7.2. ARTÍCULOS Y CAPÍTULOS DE LIBROS

Como se ha resaltado al presentar su bibliografía, se mencionan en este elenco aquellos artículos o capítulos de libros que fueron escritos como contribuciones filosóficas propiamente dichas, de modo que se omiten los textos con carácter introductorio o que sintetizan de manera sucinta aportaciones realizadas en otros trabajos.[72] Gran parte de sus artículos —más de 20— han sido luego compilados en su libro *Causality and Explanation*. El orden de presentación aquí —al igual que en el epígrafe anterior— da preferencia a la fecha de la versión original, en lugar de dar prioridad al primer apellido de los firmantes.

70. Inicialmente apareció como contribución al volumen 13 de los *Minnesota Studies in the Philosophy of Science*: SALMON, W. C., «Four Decades of Scientific Explanation», en KITCHER, PH. y SALMON, W. C. (eds.), *Scientific Explanation*, University of Minnesota Press, Minneapolis, 1989, pp. 3-219. Como libro ha sido traducido al italiano.

71. Este libro, escrito en colaboración con los profesores del Departamento de *History and Philosophy of Science* de la Universidad de Pittsburgh, ha sido traducido al griego.

72. El propio Salmon ha realizado anotaciones acerca de sus trabajos; cfr. SALMON, W. C., «Publications: An Annotated Bibliography», en FETZER, J. H. (ed.), *Probability and Causality*, Reidel, Dordrecht, 1988, pp. 271-336.

SALMON, W. C., «A Modern Analysis of the Design Argument», *Research Studies of the State College of Washington*, v. 19, n. 4, (1951), pp. 207-220.

SALMON, W. C., «The Frequency Interpretation and Antecedent Probabilities», *Philosophical Studies*, v. 4, n. 3, (1953), pp. 44-48.

SALMON, W. C., «The Uniformity of Nature», *Philosophy and Phenomenological Research*, v. 14, n. 1, (1953), pp. 39-48.

SALMON, W. C., «The Short Run», *Philosophy of Science*, v. 22, n. 3, (1955), pp. 214-221.

SALMON, W. C., «Regular Rules of Induction», *Philosophical Review*, v. 65, n. 3, (1956), pp. 385-388.

SALMON, W. C., «Should We Attempt to Justify Induction?», *Philosophical Studies*, v. 8, n. 3, (1957), pp. 33-38.

SALMON, W. C., «The Predictive Inference», *Philosophy of Science*, v. 24, n. 2, (1957), pp. 180-190.

SALMON, W. C. y NAKHNIKIAN, G., «'Exists' as a Predicate», *Philosophical Review*, v. 66, n. 4, (1957), pp. 535-542.

SALMON, W. C., «Psychoanalytic Theory and Evidence», en HOOK, S. (ed.), *Psychoanalysis, Scientific Method, and Philosophy*, New York University Press, N. York, 1959, pp. 252-267.

SALMON, W. C., «Barker's Theory of the Absolute», *Philosophical Studies*, v. 10, n. 4, (1959), pp. 50-53.

SALMON, W. C., «Vindication of Induction», en FEIGL, H. y GROVER M. (eds.), *Current Issues in the Philosophy of Science*, Holt, Rinehart and Winston, N. York, 1961, pp. 245-256.

SALMON, W. C., «On Vindicating Induction», en KYBURG, H. E. Jr. y NAGEL, E. (eds.), *Induction: Some Current Issues*, Wesleyan University Press, Middletown, Conn., 1963, pp. 27-41.

SALMON, W. C., «Inductive Inference», en BAUMRIN, B. H. (ed.), *Philosophy of Science: The Delaware Seminar*, vol. II, J. Wiley, N. York, 1963, pp. 341-370.

SALMON, W. C., «What Happens in the Long Run?», *Philosophical Review*, v. 74, n. 3, (1965), pp. 373-378.

SALMON, W. C., «The Concept of Inductive Evidence», *American Philosophical Quarterly*, v. 2, n. 4, (1965), pp. 1-6.

SALMON, W. C., «The Status of Prior Probabilities in Statistical Explanation», *Philosophy of Science*, v. 33, n. 2, (1965), pp. 137-146.

SALMON, W. C., BALKER, S. F. y KYBURG, H. E. Jr., «Symposium on Inductive Evidence», *American Philosophical Quarterly*, v. 2, (1965), pp. 265-280.

SALMON, W. C., «Consistency, Transitivity, and Inductive Support», *Ratio*, v. 7, n. 2, (1965), pp. 164-169.

SALMON, W. C., «Use, Mention, and Linguistic Invariance», *Philosophical Studies*, v. 17, nn. 1-2, (1966), pp. 13-18.

SALMON, W. C., «Verifiability and Logic», en FEYERABEND, P. K. y MAXWELL, G. (eds.), *Mind, Matter, and Method*, University of Minnesota Press, Minneapolis, 1966, pp. 347-376.

SALMON, W. C., «Carnap's Inductive Logic», *Journal of Philosophy*, v. 64, n. 21, (1967), pp. 725-739.

SALMON, W. C., «The Justification of Inductive Rules of Inference», en LAKATOS, I. (ed.), *The Problem of Inductive Logic*, North-Holland, Amsterdam, 1968, pp. 24-43.

SALMON, W. C., «Inquiries into the Foundations of Science», en ARM, D. L. (ed.), *Vistas in Science*, University of New Mexico Press, Alburquerque, 1968, pp. 1-24.

SALMON, W. C., «Partial Entailment as a Basis for Inductive Logic», en RESCHER, N. (ed.), *Essays in Honor of Carl G. Hempel*, Reidel, Dordrecht, 1969, pp. 47-82.

SALMON, W. C., «The Conventionality of Simultaneity», *Philosophy of Science*, v. 36, n. 1, (1969), pp. 44-63.

SALMON, W. C., «Statistical Explanation», en COLODNY, R. G. (ed.), *The Nature and Function of Scientific Theories*, University of Pittsburgh Press, Pittsburgh, 1970, pp. 173-232.[73]

SALMON, W. C., «Bayes's Theorem and the History of Science», en STUEWER, R. H. (ed.), *Minnesota Studies in the Philosophy of Science*, vol. V, University of Minnesota Press, Minneapolis, 1970, pp. 68-86.

SALMON, W. C., «Determinism and Indeterminism in Modern Science», en FEINBERG, J. (ed.), *Reason and Responsibility*, 2.ª ed., Dickenson, Encino, CA, 1971.

SALMON, W. C., «Memory and Perception in *Human Knowledge*», en NAKHNIKIAN, G. (ed.), *Bertrand Russell's Philosophy*, Duckworth, Londres, 1974, pp. 139-167.

SALMON, W. C., «Russell on Scientific Inference *or* Will the Real Deductivist Please Stand Up?», en NAKHNIKIAN, G. (ed.), *Bertrand Russell's Philosophy*, Duckworth, Londres, 1974, pp. 183-208.

SALMON, W. C., «An Encounter with David Hume», en FEINBERG, J. (ed.), *Reason and Responsibility*, 3.ª ed., Dickenson, Encino, CA, 1975, pp. 190-208.

SALMON, W. C., «Confirmation and Relevance», en MAXWELL, G. y ANDERSON, R. M. Jr. (eds.), *Minnesota Studies in the Philosophy*

73. Fue reimpreso en 1971 en su libro *Statistical Explanation and Statistical Relevance*, pp. 29-87.

of Science, vol. VI, University of Minnesota Press, Minneapolis, 1975, pp. 3-36.
SALMON, W. C., «Theoretical Explanation», en KÖRNER, S. (ed.), *Explanation*, B. Blackwell, Oxford, 1975, pp. 118-145.
SALMON, W. C., «Clocks and Simultaneity in Special Relativity *or* Which Twin Has the Timex?», en MACHAMER, P. K. y TURNBULL, R. G. (eds.), *Motion and Time, Space and Matter*, Ohio State University Press, Columbus, 1976, pp. 508-545.
SALMON, W. C., «Laws, Modalities, and Counterfactuals», *Synthese*, v. 35, (1977), pp. 191-230.
SALMON, W. C., «A Third Dogma of Empiricism», en BUTTS, R. y HINTIKKA, J. (eds.), *Basic Problems in Methodology and Linguistics*, Reidel, Dordrecht, 1977, pp. 149-166.
SALMON, W. C., «The Philosophy of Hans Reichenbach», *Synthese*, v. 34, n. 1, (1977), pp. 5-88.
SALMON, W. C., «Hempel's Conception of Inductive Inference in Inductive-Statistical Explanation», *Philosophy of Science*, v. 44, n. 2, (1977), pp. 180-185.
SALMON, W. C., «Indeterminism and Epistemic Relativization», *Philosophy of Science*, v. 44, (1977), pp. 199-202.
SALMON, W. C., «An 'At-At' Theory of Causal Influence», *Philosophy of Science*, v. 44, n. 2, (1977), pp. 215-224.
SALMON, W. C., «The Curvature of Physical Space», en EARMAN, J. S., GLYMOUR, C. N. y STACHEL, J. (eds.), *Minnesota Studies in the Philosophy of Science*, vol. VIII, University of Minnesota Press, Minneapolis, 1977.
SALMON, W. C., «The Philosophical Significance of the One-way Speed of Light», *Nous*, v. 9, n. 3, (1977), pp. 353-392.
SALMON, W. C., «An Ontological Argument for the Existence of the Null Set», en GARDNER, M., *Mathematical Magic Show*, Alfred A. Knopf, N. York, 1977, pp. 32-33.
SALMON, W. C., «Objectively Homogeneous Reference Classes», *Synthese*, v. 36, n. 4, (1977), pp. 399-414.
SALMON, W. C., «Unfinished Business: The Problem of Induction», *Philosophical Studies*, v. 33, (1978), pp. 1-19.
SALMON, W. C., «Why Ask, 'Why?'?—An Inquiry Concerning Scientific Explanation», *Proceedings and Addresses of the American Philosophical Association*, v. 51, n. 6, (1978), pp. 683-705.
SALMON, M. H. y SALMON, W. C., «Alternative Models of Scientific Explanation», *American Anthropologist*, v. 81, n. 1, (1979), pp. 61-74.
SALMON, W. C., «A Philosopher in a Physics Course», *Teaching Philosophy*, v. 2, n. 2, (1979), pp. 139-146.

SALMON, W. C., «Propensities: A Discussion Review», *Erkenntnis*, v. 14, (1979), pp. 183-216.

SALMON, W. C., «Informal Analytic Approaches to the Philosophy of Science», en ASQUITH, P. y KYBURG, H. E. Jr. (eds.), *Current Research in Philosophy of Science*, Philosophy of Science Association, East Lansing, Mich., 1979, pp. 3-15.

SALMON, W. C., «Probabilistic Causality», *Pacific Philosophical Quarterly*, v. 61, nn. 1-2, (1980), pp. 50-74.

SALMON, W. C., «John Venn's *Logic of Chance*», en HINTIKKA, J., GRUENDER, D. y AGAZZI, E. (eds.), *Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science*, Reidel, Dordrecht, 1980, pp. 125-138.

SALMON, W. C., «Robert Leslie Ellis and the Frequency Theory», en HINTIKKA, J., GRUENDER, D. y AGAZZI, E. (eds.), *Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science*, Reidel, Dordrecht, 1980, pp. 139-143.

SALMON, W. C., «Causality: Production and Propagation», *Proceedings of the Philosophy of Science Association*, v. 2, (1980), pp. 49-69.

SALMON, W. C., «Probabilistic Causality», *Pacific Philosophical Quarterly*, v. 61, (1980), pp. 50-74.

SALMON, W. C., «Causality: Production and Propagation», en ASQUITH, P. D. y GIERE, R. N. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing, Michigan, 1981, pp. 49-69.

SALMON, W. C., «Rational Prediction», *British Journal for the Philosophy of Science*, v. 32, n. 2, (1981), pp. 115-125.[74]

SALMON, W. C., «In Praise of Relevance», *Teaching Philosophy*, v. 4, nn. 3-4, (1981), pp. 261-275.

SALMON, W. C., «Comets, Pollen, and Dreams: Some Reflections on Scientific Explanation», en MCLAUGHLIN, R. (ed.), *What? Where? When? Why?*, Reidel, Dordrecht, 1982, pp. 261-275.

SALMON, W. C., «Further Reflections», en MCLAUGHLIN, R. (ed.), *What? Where? When? Why?*, Reidel, Dordrecht, 1982, pp. 231-280.

SALMON, W. C., «Causality: Production and Propagation», en ASQUITH, P. D. ET AL. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing, Michigan, 1982, pp. 49-69.

74. Reimpreso como SALMON, W. C., «Rational Prediction», en GRÜNBAUM, A. y SALMON, W. C. (eds.), *The Limitations of Deductivism*, University of California Press, Berkeley, 1988, pp. 47-60. En este libro ofrece una amplia introducción, pp. 1-18.

SALMON, W. C., «Causality in Archaeological Explanation», en RENFREW, C. ET AL. (eds.), *Theory and Explanation in Archaeology*, Academic Press, N. York, 1982, pp. 45-55.

SALMON, W. C., «Probabilistic Explanation: Introduction», en ASQUITH, P. y NICKLES, TH. (eds.), *PSA 1982*, Philosophy of Science Association, East Lansing, Michigan, 1983, pp. 179-180.

SALMON, W. C., «Carl G. Hempel on the Rationality of Science», *Journal of Philosophy*, v. 80, n. 10, (1983), pp. 555-562.

SALMON, W. C., «Scientific Explanation: Three Basic Conceptions», *Proceedings of the Philosophy of Science Association*, v. 2, (1984), pp. 293-305.

SALMON, W. C., «Empiricism: The Key Question», en RESCHER, N. (ed.), *The Heritage of Logical Positivism*, University Press of America, Lanham, MD, 1985, pp. 1-21.

SALMON, W. C., «Conflicting Conceptions of Scientific Explanation», *Journal of Philosophy*, v. 82, n. 11, (1985), pp. 651-654.

SALMON, W. C., «Scientific Explanation: Three Basic Conceptions», en ASQUITH, P. y KITCHER, PH. (eds.), *PSA 1984*, Philosophy of Science Association, East Lansing, Michigan, 1986, pp. 293-305.

KITCHER, PH. y SALMON, W. C., «Van Fraassen on Explanation», *Journal of Philosophy*, v. 84, (1987), pp. 315-330.[75]

SALMON, W. C., «Statistical Explanation and Causality», en PITT, J. (ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988, pp. 75-118.[76]

SALMON, W. C., «Dynamic Rationality: Propensity, Probability, and Credence», en FETZER, J. H. (ed.), *Probability and Causality*, Reidel, Dordrecht, 1988, pp. 3-40.

SALMON, W. C., «Deductivism Visited and Revisited», en GRÜNBAUM, A. y SALMON, W. C. (eds.), *The Limitations of Deductivism*, University of California Press, Berkeley, 1988, pp. 95-127.

SALMON, W. C., «Intuitions: Good and Not-So-Good», en HARPER, W. y SKYRMS, B. (eds.), *Causation, Chance, and Credence*, Reidel, Dordrecht, 1988, pp. 51-71.

SALMON, W. C., «Statistical Explanation and its Models», en BRODY, B. A. y GRANDY, R. E. (eds.), *Readings in the Philosophy of Science*, 2.ª ed., Prentice-Hall, Englewood Cliffs, NJ, 1989, pp. 167-83.[77]

75. Este artículo fue reimpreso en RUBEN, D. H. (ed.), *Explanation*, Oxford University Press, Oxford, 1993, pp. 78-112.

76. Contiene un extracto de *Statistical Explanation and Statistical Relevance* y los capítulos 5 y 6 de *Scientific Explanation and the Causal Structure of the World.*

77. Reimpresión del capítulo 2 de *Scientific Explanation and the Causal Structure of the World.*

SALMON, W. C., «Causal Propensities: Statistical Causality vs. Aleatory Causality», *Topoi*, v. 9, (1990), pp. 95-100.

SALMON, W. C., «Rationality and Objectivity in Science, *or* Tom Kuhn Meets Tom Bayes», en SAVAGE, C. W. (ed.), *Scientific Theories*, University of Minnesota Press, Minneapolis, 1990, pp. 175-204.

SALMON, W. C., «Scientific Explanation: Causation *and* Unification», *Crítica*, v. 22, (1990), pp. 3-21.

SALMON, W. C., «Philosophy and the Rise of Modern Science», *Teaching Philosophy*, v. 13, (1990), pp. 233-239.

SALMON, W. C., «Propensioni causali: causalità statistica e causalità aleatoria», *Rivista di filosofia*, v. 81, (1990), pp. 167-180.

SALMON, W. C., «Hans Reichenbach's Vindication of Induction», *Erkenntnis*, v. 35, (1991), pp. 99-122.

SALMON, W. C., «Scientific Explanation», en SALMON, M. ET AL., *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs, NJ, 1992, pp. 7-41.

EARMAN, J. y SALMON, W. C., «The Confirmation of Scientific Hypotheses», en SALMON, M. ET AL., *Introduction to the Philosophy of Science*, Prentice Hall, Englewood Cliffs, NJ, 1992, pp. 42-103.

SALMON, W. C., «Explanation in Archaeology: An Update», en EMBREE, L. (ed.), *Metaarchaeology*, Kluwer, Dordrecht, 1992, pp. 243-253.

SALMON, W. C., «The Value of Scientific Understanding», *Philosophica*, v. 5, (1993), pp. 9-19.

SALMON, W. C., «Scientific Explanation and the Causal Structure of the World», en RUBEN, D. H. (ed.), *Explanation*, Oxford University Press, Oxford, 1993, pp. 72-112.[78]

SALMON, W. C., «On the Alleged Temporal Anisotropy of Explanation», en EARMAN, J., JANIS, A., MASSEY, G. y RESCHER, N. (eds.), *Philosophical Problems of the Internal and External Worlds*, University of Pittsburgh Press, Pittsburgh, 1993, pp. 229-248.

SALMON, W. C., «Causality without Counterfactuals», *Philosophy of Science*, v. 61, (1994), pp. 297-312.

SALMON, W. C., «Carnap, Hempel, and Reichenbach on Scientific Realism», en SALMON, W. C. y WOLTERS, G. (eds.), *Logic, Language, and the Structure of Scientific Theories*, University of Pittsburgh Press, Pittsburgh, 1994, pp. 237-254.

78. Recoge una selección procedente de *Scientific Explanation and the Causal Structure of the World.*

SALMON, W. C., «La comprensión científica en el siglo veinte», *Revista Latinoamericana de Filosofía*, v. 21, n. 1, (1995), pp. 3-21.

SALMON, W. C., «La causalità», *Iride*, v. 8, n. 14, (1995), pp. 187-196.

SALMON, W. C., «Causality and Explanation: A Reply to Two Critiques», *Philosophy of Science*, v. 64, (1997), pp. 461-477.

SALMON, W. C., «A Contemporary Look at Zeno's Paradoxes», en INWAGEN, P. VAN y ZIMMERMAN, D. W. (eds.), *Metaphysics: The Big Questions*, Blackwell, Oxford, 1998, pp. 129-149.[79]

SALMON, W. C., «Scientific Explanation: How We Got from There to Here», en KLEMKE, E. D. ET AL. (eds.), *Introductory Readings in the Philosophy of Science*, Promethius Books, Amherst, NY, 1998, pp. 241-263.[80]

SALMON, W. C., «The Problem of Induction», en PERRY, J. y BRATMAN, M. (eds.), *Introduction to Philosophy*, 3.ª ed., Oxford University Press, N. York, 1999, pp. 230-251.[81]

SALMON, W. C., «Der Schluß auf die beste Erklärung», en LUEKEN, G.-L. y STEKELER-WEITHOFER, P. (eds.), *Argumentation*, Felix Meiner, Hamburgo, 1999, pp. 115-139.[82]

SALMON, W. C., «The Spirit of Logical Empiricism: Carl G. Hempel's Role in Twentieth-Century Philosophy of Science», *Philosophy of Science*, v. 66, (1999), pp. 333-350.

SALMON, W. C., «Ornithology in a Cubical World: Reichenbach on Scientific Realism», en GREENBERGER, D. ET AL. (eds.), *Epistemological and Experimental Perspectives on Quantum Physics*, Kluwer, Dordrecht, 1999, pp. 305-315.

SALMON, W. C., «Scientific Understanding in the Twentieth Century», en REDEI, M. y STÖLTZNER, M. (eds.), *John von Neumann and the Foundations of Quantum Physics*, Kluwer, Dordrecht, 2001, pp. 289-304.

SALMON, W. C., «Explaining Things Probabilistically», *The Monist*, v. 84, n. 2, (2001), pp. 208-217.

SALMON, W. C., «Explicación causal frente a no causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 97-115.

SALMON, W. C., «Estructura de la explicación causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 141-159.

79. Constituye el capítulo 2 de *Space, Time, and Motion*.
80. Reimpreso de *Causality and Explanation*, pp. 302-319.
81. Basado en *Foundations of Scientific Inference*.
82. Traducción alemana de la conferencia «Inference to the Best Explanation», dada en Leipzig en 1996.

SALMON, W. C., «Explanation and Confirmation: Inference to the Best Explanation», en HON, G. y RAKOVER, S. S. (eds.), *Explanation: Theoretical Approaches and Applications*, Kluwer, Dordrecht, en prensa.

7.3. PRINCIPALES PUBLICACIONES SOBRE SU PENSAMIENTO

Respecto de las publicaciones sobre W. C. Salmon cabe resaltar tres compilaciones: la primera tiene su origen en una visita que realizó a Australia (septiembre-diciembre de 1978); la segunda es una recopilación de trabajos en torno a los temas centrales de su Pensamiento; y la tercera corresponde a un Simposio en honor suyo y de Merrilee Salmon, organizado por el Centro de Historia y Filosofía de la Ciencia de Florencia en mayo de 1996. Las referencias de esos volúmenes, según el orden de presentación habitual, son las siguientes:

FETZER, J. H. (ed.), *Probability and Causality: Essays in Honor of Wesley C. Salmon*, Reidel, Dordrecht, 1988.

GALAVOTTI, M. C. y PAGNINI, A. (eds.), *Experience, Reality, and Scientific Explanation: Essays in Honor of Merrilee and Wesley Salmon*, Kluwer, Dordrecht, 1999.

MCLAUGHLIN, R. (ed.), *What? Where? When? Why? - Essays on Induction, Space and Time, Explanation*, Reidel, Dordrecht, 1982.

8. **Bibliografía sobre explicación científica**

Durante varias décadas, el estudio filosófico-metodológico de la explicación científica ha tenido particular protagonismo, por lo que un elenco bibliográfico exhaustivo acerca de este tema excedería con mucho las páginas disponibles en este libro. Hay, además, un excelente libro que recoge una amplia bibliografía del largo periodo que analiza (1948-1987), donde se añaden asimismo una serie de trabajos anteriores a esa etapa y también alguno posterior: es el volumen de Wesley C. Salmon *Four Decades of Scientific Explanation*, en sus páginas 196-219.

Tras recomendar esa detallada relación bibliográfica, y después de citar los trabajos del principal especialista contemporáneo —W. C. Salmon—, tanto solo como en colaboración con Philip Kitcher, se seleccionan aquí algunos estudios de las décadas cubiertas por ese volumen y enumeran también diversas contribuciones posterio-

res a ese libro. Esas publicaciones se orientan, básicamente, en una de estas dos direcciones: i) el estudio de la naturaleza de la «explicación científica», en cuanto tal, o de alguno de los tipos de explicaciones que se dan en la Ciencia (en especial, la causal); o ii) el análisis de las relaciones con la predicción científica, bien sea en términos de simetría-asimetría o bien al hilo del problema de cuál de las dos —explicación o predicción— posee mayor peso específico para la Ciencia.

ACHINSTEIN, P., *Law and Explanation: An Essay in the Philosophy of Science*, Clarendon Press, Oxford, 1971.

ACHINSTEIN, P., *The Nature of Explanation*, Oxford University Press, Oxford, 1983.

ACHINSTEIN, P., «Explanation v. Prediction: Which Carries More Weight?», en HULL, D., FORBES, M. y BURIAN, R. M. (eds.), *PSA 1994*, Philosophy of Science Association, East Lansing, Michigan, vol. 2, 1995, pp. 156-164.

ANSCOMBE, G. E. M., *Causality and Determination*, Cambridge University Press, Cambridge, 1971.

BRUSH, S. G., «Dynamics of Theory Change: The Role of Predictions», en HULL, D., FORBES, M. y BURIAN, R. M. (eds.), *PSA 1994*, Philosophy of Science Association, East Lansing, Michigan, vol. 2, 1995, pp. 133-145.

DOWE, PH., «Causality and Explanation», *British Journal for the Philosophy of Science*, v. 51, (2000), pp. 165-174.

DRAY, W., *Laws and Explanation in History*, Clarendon Press, Oxford, 1957 (reimp. en 1964).

EARMAN, J. (ed.), *Inference, Explanation and Other Frustrations. Essays in the Philosophy of Science*, University of California Press, Berkeley, 1992.

FETZER, J. H. (ed.), *Science, Explanation, and Rationality: The Philosophy of Carl G. Hempel*, Oxford University Press, N. York, 2000.

FRAASSEN, B. C. VAN, «The Pragmatics of Explanation», *American Philosophical Quarterly*, v. 14, (1977), pp. 143-150.

FRAASSEN, B. C. VAN, *The Scientific Image*, Clarendon Press, Oxford, 1980.

FRAASSEN, B. C. VAN, «The Pragmatic Theory of Explanation», en PITT, J. C. (ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988, pp. 136-155.

FRIEDMAN, M., «Explanation and Scientific Understanding», *Journal of Philosophy*, v. 71, (1974), pp. 5-19. Compilado en PITT, J. C.

(ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988, pp. 188-198.

GONZÁLEZ, W. J., «Lakatos's Approach on Prediction and Novel Facts», *Theoria*, v. 16, n. 42, (2001), pp. 499-518.

GONZÁLEZ, W. J., «From *Erklären-Verstehen* to *Prediction-Understanding*: The Methodological Framework in Economics», en SINTONEN, M., YLIKOSKI, P. y MILLER, K. (eds.), *Realism in Action: Essays in the Philosophy of Social Sciences*, Kluwer, Dordrecht, 2002, pp. 33-48.

GRÜNBAUM, A., «Temporally Asymetric Principles, Parity between Explanation and Prediction, and Mechanism versus Teleology», *Philosophy of Science*, v. 29, (1962), pp. 146-170.

GRUNER, R., «Teleological and Functional Explanations», *Mind*, v. 75, (1966), pp. 516-526.

HAUSMAN, D., *Causal Asymmetries*, Cambridge University Press, Cambridge, 1998.

HEMPEL, C. G., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, N. York, 1965.

HEMPEL, C. G., *Philosophy of Natural Science*, Prentice Hall, Englewood Cliffs, NJ, 1966.

HOWSON, C., «Accomodation, Prediction and Bayesian Confirmation Theory», en FINE, A. y LEPLIN, J. (ed.), *PSA 1988*, Philosophy of Science Association, East Lansing, Michigan, vol. 2, 1989, pp. 381-392.

HOWSON, C., «Fitting your Theory to the Facts: Probably not such a Bad Thing after All», en WADE SAVAGE, C. (ed.), *Scientific Theories*, Minnesota Studies in the Philosophy of Science, Minneapolis, 1990, pp. 224-244.

JEFFREY, R. C., «Statistical Explanation vs. Statistical Inference», en RESCHER, N. ET AL. (eds.), *Essays in Honor of Carl G. Hempel*, Reidel, Dordrecht, 1969, pp. 104-113.

JOYNT, C. B. y RESCHER, N., «On Explanation in History», *Mind*, v. 68 n. 271 (1959), pp. 383-387.

KIM, J., «On the Logical Conditions of Deductive Explanation», *Philosophy of Science*, v. 30, (1963), pp. 286-291.

KITCHER, PH., «Explanatory Unification», *Philosophy of Science*, v. 48, (1981), pp. 507-531.

KOERTGE, N., «Explanation and its Problems», *British Journal for the Philosophy of Science*, v. 43, (1992), pp. 85-98.

KÖRNER, S. (ed.), *Explanation*, B. Blackwell, Oxford, 1975.

LIPTON, P., *Inference to the Best Explanation*, Routledge, Londres, 1993.

LIPTON, P. (ed.), *Theory, Evidence and Explanation*, Aldershot, Dartmouth, 1995.

LITTLE, D., *Varieties of Social Explanation*, Westview Press, Summertown, Oxford, 1991.

LLOYD, C., *Explanation in Social History*, B. Blackwell, Oxford, 1986.

MANNINEN, J. y TUOMELA, R. (eds.), *Essays on Explanation and Understanding*, Reidel, Dordrecht, 1976.

MCINTYRE, L. C., *Laws and Explanation in the Social Sciences: Defending a Science of Human Behavior*, Westwiew Press, Boulder, CO, 1996.

NAGEL, E., *The Structure of Science*, Harcourt, Brace and World, N. York, 1961.

NIINILUOTO, I., «Statistical Explanation Reconsidered», *Synthese*, v. 48, (1981), pp. 437-472.

NIINILUOTO, I., *Critical Scientific Realism*, Clarendon Press, Oxford, 1999.

PITT, J. C. (ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988.

RAILTON, P., «Probability, Explanation, and Information», *Synthese*, v. 48, (1981), pp. 233-256.

RESCHER, N., «The Stochastic Revolution and the Nature of Scientific Explanation», *Synthese*, v. 14, (1962), pp. 200-215.

RESCHER, N. y SKYRMS, B., «A Methodological Problem in the Evaluation of Explanation», *Nous*, v. 2, (1968), pp. 121-129.

RESCHER, N., *Scientific Explanation*, Free Press, N. York, 1970.

RESCHER, N., *Predicting the Future*, State University Press New York, N. York, 1998.

ROSENBERG, A., *Philosophy of Social Science*, 2.ª edición, Westview, Boulder, CO, 1995.

RUBEN, D. H., *Explaining Explanation*, Routledge, Londres, 1990.

RUBEN, D. H. (ed.), *Explanation*, Oxford University Press, Oxford, 1993.

RYAN, A. (ed.), *The Philosophy of Social Explanation*, Oxford University Press, Oxford, 1973.

SCHEFFLER, I., «Explanation, Prediction, and Abstraction», *British Journal for the Philosophy of Science*, v. 7, (1957), pp. 293-309.

SCHWARTZ, J., «Functional Explanation and Metaphysical Individualism», *Philosophy of Science*, v. 60, (1993), pp. 278-301.

SCRIVEN, M., «Explanations, Predictions and Laws», en FEIGL, H. y MAXWELL, G. (eds.), *Scientific Explanation, Space and Time*, University of Minnesota Press, Minneapolis, 1962, pp. 170-230. Compilado en PITT, J. C. (ed.), *Theories of Explanation*, Oxford University Press, N. York, 1988, pp. 51-74.

SCRIVEN, M., «Causation as Explanation», *Nous*, v. 9, (1975), pp. 3-16.

SIMON, H. A. y RESCHER, N., «Cause and Counterfactual», *Philosophy of Science*, v. 33, (1966), pp. 323-340.

SIMON, H. A., «Discovering Explanations», *Minds and Machines*, v. 8, n. 1, (1998), pp. 7-37.

SINTONEN, M., *The Pragmatics of Scientific Explanation*, Societas Philosophica Fennica, Helsinki, 1984.

SOBER, E., «Explanation and Causation», *British Journal for the Philosophy of Science*, v. 38, (1987), pp. 243-257.

STEGMÜLLER, W., «Explanation, Prediction, Scientific Systematization and Non-Explanatory Information», *Ratio*, v. 8, (1966), pp. 1-24.

SUPPES, P. y ZANOTTI, M., «When are Probabilistic Explanations Possible?», *Synthese*, v. 48, (1981), pp. 191-199.

TUOMELA, R., «Dispositions, Realism and Explanation», *Synthese*, v. 34, (1977), pp. 457-487.

TUOMELA, R., «Inductive Explanation», *Synthese*, v. 48, (1981), pp. 257-294.

TUOMELA, R., «Truth and the Best Explanation», en ESSLER, W. K., PUTNAM, H. y STEGMÜLLER, W. (eds.), *Epistemology, Methodology and Philosophy of Science*, Reidel, Dordrecht, 1985, pp. 217-299.

WILSON, F., *Explanation, Causation and Deduction*, Reidel, Dordrecht, 1985.

WOODWARD, J., «Explanation and Invariance in the Special Sciences», *British Journal for the Philosophy of Science*, v. 51, (2000), pp. 197-254.

WOODWARD, J., «Law and Explanation in Biology: Invariance is the Kind of Stability that Matters», *Philosophy of Science*, v. 68, (2001), pp. 1-20.

WRIGHT, G. H. VON, *Explanation and Understanding*, Cornell University Press, Ithaca, 1971.

WRIGHT, L., *Teleological Explanations*, University of California Press, Berkeley, 1976.

PARTE II

EVOLUCIÓN DEL PROBLEMA DE LA EXPLICACIÓN CIENTÍFICA

Capítulo 2

EXPLICACIÓN CIENTÍFICA Y COMPLEJIDAD[1]

Javier Ordóñez

1. Contexto temático

Un libro donde se trate expresamente de la explicación científica puede ser considerado por muchos como inadecuado para un historiador de la Ciencia. La razón es muy sencilla, la Historia de la Ciencia se presenta hoy como una disciplina muy plural, tributaria de una gran variedad de tradiciones, que trabaja en muchas direcciones y, en apariencia, todas ellas relacionadas con un profundo relativismo y siempre con una actitud de menosprecio hacia los problemas centrales de la Filosofía de la Ciencia del siglo pasado, es decir, del siglo XX. El historicismo ha sido un huracán intelectual que pretende haber arrasado antiguas convicciones acerca de qué se puede saber sobre la construcción de la Ciencia arrumbándolas como si sólo fueran viejas preguntas producto de unas modas de generaciones de filósofos del pasado cercano. Mi punto de vista acerca de estas cuestiones no se identifica con este historicismo tan disolvente: defiendo otras formas de hacer la Historia de la Ciencia. Efectivamente, pienso que el carácter dinámico del conocimiento científico es un problema intelectual muy relevante, que debe estar presente en las reflexiones sobre la explicación científica y en otros problemas clásicos como la cuestión del realismo. Considero, en definitiva, que es una forma posible de hacer Filosofía de la Ciencia o, simplemente, Filosofía.

Desde mi punto de vista, la Historia no sirve exclusivamente para introducir una sana dosis de escepticismo en el triunfalismo de los científicos, como pretenden los que defienden su reducción a

1. Agradezco a Wenceslao J. González las sugerencias para la ponencia inicial y la versión final de este texto.

pura Sociología. Tampoco debe servir para cultivar la esperanza de poder reconstruir la Ciencia con una simplicidad académica que nos permita hablar de ella como si fuera el resumen de una actividad racional siempre guiada por motivos exclusivamente teóricos y traducibles en todo instante a procesos racionales. No conviene a mi juicio representar la Ciencia como una actividad truncada por los resúmenes que nos ofrecen los científicos en forma de tratados sumarios, en forma de manuales, aptos sin duda para la enseñanza universitaria pero incapaces de dar la menor indicación de los procesos de construcción científica. Esta perspectiva sincrónica sólo resulta alentadora a quienes pretenden mostrar que el despliegue científico es un proceso racional inmediatamente reconocible como ocurre con los discípulos que siguieron las ideas de Lakatos. Se podría decir que en el curso de sus trabajos han encontrado todo lo que han puesto en su búsqueda, pero sólo lo que han puesto. Sus seguidores han hecho bueno el *dictum* de su maestro,[2] cuya paráfrasis podría ser: «si la Historia es ciega, su racionalidad es la Filosofía que debe iluminar el mundo de la Ciencia». Pero la Historia puede que no sea tan ciega ni su Filosofía tan iluminadora.

Pero hay al menos otra forma de abordar la Historia de la Ciencia que permita dialogar con la Filosofía, sea de la Ciencia o no. Su punto de partida sería no asustarse por la complejidad que manifiesta la Ciencia durante el proceso de su producción. No quiero decir con esto que debamos tomar sin restricción lo que las fuentes históricas nos proporcionan, sino simplemente reconocer que el primer sistema complejo con el que se topa la Ciencia es el de su misma Historia. Si lo hacemos así, reconoceremos que es necesario tomar en consideración no sólo la producción del conocimiento sino, tanto los problemas metodológicos que se tomaron en cuenta para producirlo, como las cuestiones filosóficas que discutieron aquellos científicos que trataron de construir teorías científicas y que trabajaron en los contextos intelectuales necesarios para ello. Cualquier aproximación histórica a la Ciencia muestra que los debates y trabajos que la produjeron no se limitaron a la cuestión de elección entre teorías rivales. Muchas veces, más de las que habitualmente se quiere reconocer, tuvieron lugar amargas polémicas entre concepciones metodológicas rivales que abrieron el paso a las novedades científicas más radicales. Una de ellas va

2. Me refiero a la afirmación «philosophy of science without history is empty; history of science without philosophy of science is blind», Lakatos, I. (1978), p. 102. Según algunos, es una afirmación muy comprensible e, incluso, fuera de toda discusión.

a ser el tema fundamental de esta conferencia: la que permitió discutir por primera vez sobre la posibilidad de hablar de sistemas complejos en la explicación científica. Desde mi punto de vista, tratar con más frecuencia estos problemas acercaría la Historia a la Filosofía de la Ciencia tradicional y a la misma Ciencia, se considere o no en su contexto histórico.

2. Afinidad entre Filosofía y Ciencia

Habitualmente se insiste en que la Historia de la Ciencia es mala compañera de la «Concepción heredada» y, desde luego, de los problemas relacionados con la explicación científica. Asimismo, se destaca que el contexto de justificación, ámbito donde se desarrollan los tópicos más relevantes de la «Concepción heredada», no necesita ningún recurso a la Historia de cómo se construyó la Ciencia. Como contrapartida, sus críticos señalan que en el contexto en el que ellos trabajan no se tratan problemas científicos, sino exclusivamente filosóficos alejados de cualquier forma de Ciencia real, tanto del presente como del pasado. Y esa crítica se extiende a cualquier filósofo de la Ciencia que no tome la Historia como fuente de información desde un punto de vista explícito. No pretendo, como se comprenderá, zanjar la polémica a favor de uno u otro punto de vista.

Como historiador, doy un valor especial al estudio de la formación de las Ciencias, que no se reduce únicamente a la formación de las teorías científicas. Pero creo que los mismos filósofos de la Ciencia, por mucho que sus intereses intelectuales estén alejados de la Historia, reproducen comportamientos que ya se dieron en el seno de las discusiones entre científicos. Es decir, que en la Filosofía de la Ciencia se vuelven a tratar cuestiones parecidas a las del pasado; con nuevas herramientas, sin duda y, por supuesto, con un refinamiento mayor. En este sentido, algunos historiadores de la Ciencia tenemos la sensación de asistir a la reproducción benigna del mito de Sísifo. Problemas inevitables aparecen una y otra vez en el horizonte de la Filosofía, y serían una maldición, si no fuera un castigo gozoso volverse a plantear permanentemente las mismas cuestiones que afectan no sólo a nuestra concepción de la Ciencia sino a nuestra interpretación del mundo.

En el presente volumen se aborda expresamente la explicación científica y sus problemas anejos. En cualquiera de *las versiones* que ha estudiado el profesor Wesley C. Salmon se manifiesta la preten-

sión de reproducir procesos metodológicos muy simples en un intento de reducir la Ciencia a su expresión metodológica más elemental y estudiar a partir de ella las condiciones en las que se pudiera expresar la causalidad, debatiendo siempre el problema dentro de la dualidad realismo-empirismo. Ante este intento, podría decirse que durante siglos la Ciencia también ha pretendido reducir el conocimiento a la simplicidad más pura. La Historia de ese proceso, que se denomina sumariamente como «revolución científica», también puede contarse como la búsqueda de un elemento de simplicidad que pudiera resumir la complejidad de los procesos naturales. De hecho, puede afirmarse sin temor a exagerar que la nueva Ciencia mostró poco interés por las complejidades, más propias de la época clásica o antigua, y que se centró casi exclusivamente en la Metodología cartesiana que pretendía resolver las dificultades por división en elementos simples.

De Descartes a Whitehead, la Ciencia pareció obedecer a los mismos dictados. Este último filósofo aseguraba que la Ciencia debía buscar las explicaciones más simples de los fenómenos más complejos. Ambos, y todos los personajes intermedios, sintonizaban con el nominalismo de un principio de Occam que en términos contemporáneos podría expresarse de la siguiente manera: de todas las expresiones matemáticas que den cuenta de un fenómeno, la más simple es la «verdadera».

Según esto, la vocación de la Ciencia moderna consistió en sustituir la complejidad de los fenómenos naturales que se ofrecen a nuestros sentidos por la simplicidad de procesos elementales. Fue la sustitución de una «Física» de la Naturaleza, entendida al modo antiguo de la tradición aristotélica, por una «Mecánica» teórica concebida en términos barrocos. Parafraseando a Koyré, se podría decir que la distinción entre sublunar-supralunar, propia de la concepción antigua de la Naturaleza, fue sustituida por la división entre lo cuantitativo —o propio de la geometría— y lo cualitativo, o el mundo de la complejidad en el que naufragaba cualquier intento de explicación.[3]

La mejor expresión de esa ruptura se encuentra en la introducción del corpuscularismo cartesiano, primero, y del atomismo newtoniano, después. Tanto Descartes como Newton impusieron las condiciones para representar fenómenos complejos por medio de sistemas muy simples donde se pudiera estudiar con precisión lo que podríamos denominar como «acontecimiento primordial» del

3. Un análisis interesante del pensamiento de A. Koyré se encuentra en la introducción de Carlos Solís a Koyré, A. (1994), pp. 9-43.

mundo (en este primer periodo no se hablaba de «fenómenos», ya que éstos no se revelaban a los sentidos). Para Descartes, ese acontecimiento era el empujón entre corpúsculos ya que no aceptaba la existencia del vacío. En el caso de Newton, se elegía el choque entre átomos como un enfoque más clásico de una indudable tradición discontinuista. En ambos casos se reconocían dos sistemas mecánicos donde se daban siempre dos condiciones básicas y una suposición implícita. Las dos condiciones eran:

Primera, que se pudiera identificar cuál de los corpúsculos o los átomos era el primero y cuál el segundo. *Segunda,* que se transmitiera algo. Podía ser la velocidad, la cantidad de movimiento, la *vis viva* o cualquiera otra magnitud mecánica. Además, *la condición añadida* consistía en que nunca se podía hablar del agente exterior al proceso. El movimiento de un corpúsculo o de un átomo siempre era debido a otro corpúsculo o a otro átomo, en un recurso que la interpretación siempre podía convertir en indefinido, si no en infinito.

La difusión de las ideas de Newton es suficientemente conocida como para no insistir en ella. Pero su éxito entre sus colegas residió en gran medida en que ese sistema mecánico simple del choque podía ser refinado en términos más abstractos de *interacción* que no siempre requería el contacto físico. Su Mecánica admitía que los cuerpos fueran simplificados hasta el extremo de ser tratados como puntos, y sus acciones como interacciones, que en «nuestro» mundo podrían deberse a una causa de movimiento primordial llamada fuerza gravitatoria. Todo esto, además se podía expresar con mayor o menor trabajo de forma analítica. Las reconstrucciones posteriores de la Mecánica muestran la potencia del modelo newtoniano, tanto por lo que toma, como por lo que abandona a modo de lastre epistemológico innecesario.

Los trabajos de los matemáticos ilustrados son muy elocuentes en este sentido. Se supone la existencia de un sistema mecánico primordial basado siempre en una urdimbre corpuscular o atomista que forma la base de la marcha de la construcción del análisis matemático. Y no es una casualidad que ocurriera esto. Mecánica y análisis eran isomorfos, al menos en ese momento.[4] Pero precisamente ese momento es el que ha quedado congelado en la memoria de muchos filósofos de la Ciencia contemporáneos. Y no porque desprecien la Historia, sino porque simplemente eligen exclusiva-

4. Sobre la relación entre el análisis barroco y la Mecánica puede consultarse en Blay, M. (1992), especialmente pp. 277-321, y en Truesdell, C. (1968), pp. 223-251

mente un momento de ella, aunque sin duda, por razones muy poderosas. A partir de ahí parecen recorrer el mismo itinerario que los científicos, lo reconozcan o no.

En este punto, un inciso puede ser de utilidad. En general se suele destacar que la Filosofía de la Ciencia de primera hora está impregnada de fisicalismo que, según se dice, contamina toda su tarea posterior. En mi opinión, es demasiado ingenuo hablar de fisicalismo en el análisis filosófico de las Ciencias de la Naturaleza. En rigor, ni siquiera en la Física ha habido un fisicalismo porque en su seno sí se ha producido un debate epistemológico sobre la pertinencia de tomar en cuenta sistemas complejos que nunca se ha visto reflejado en la Filosofía. Más bien ha habido un mecanicismo. Es decir, la Filosofía de la Ciencia, especialmente la que se preocupa por la explicación científica, ha recogido el debate sobre la importancia de la Mecánica como término de reducción de todo el resto de las teorías físicas y, de esta forma, se ha situado en uno de los lados de la polémica.

Pero, sin duda, es un lado ilustre que usa sus argumentos con acierto. Si se siguen los ejemplos más utilizados para hablar de la finalidad de la Ciencia en un autor tan incisivo como van Fraassen, cuando contrapone el realismo al empirismo crítico en su obra *The Scientific Image*,[5] encontramos que la mayor parte de ellos están extraídos de las diferentes etapas de la Mecánica o de autores que creían firmemente en la posibilidad de reducción de la Física a la Mecánica, aunque elaboraran teorías que trataran de fenomenologías tales como el electromagnetismo. Aunque después aborde otros temas de Física, la puerta de entrada es la Mecánica que constituye como un elemento epistémico imprescindible para cualquier otra incursión científica posterior. Parece que hablar de mundo externo requiere hablar de objetos que tengan la consistencia de los físicos, y dentro de ellos, de los mecánicos.

Veamos las dos definiciones que van Fraassen ofrece en esa obra acerca de cuál es el propósito de la Ciencia para los *realistas* y para los *empiristas constructivos*:

Los primeros dirían que «la finalidad de la Ciencia es darnos en sus teorías una verdad literal: la Historia de lo que el mundo es; y la aceptación de una teoría científica incluye la creencia de que es verdadera».[6] Los segundos, por el contrario, considerarían que «la finalidad de la Ciencia consiste en dar teorías adecuadas empírica-

5. Fraassen, B. van (1980).
6. Fraassen, B. van (1980), p. 8

mente; y aceptar una teoría significa que se cree sólo que es empíricamente adecuada».[7]

Dejando al margen las consecuencias de tales definiciones y los debates que implican, que están suficientemente recogidos en la obra de B. van Fraassen, se puede destacar que esa forma de ver la Ciencia restringe su actividad a un ámbito muy reducido. La reduce sólo a la producción de teorías como si éstas se pudieran distinguir del resto de la actividad científica por medio de un hipotético escalpelo metodológico. Una vez más, esta perspectiva filosófica va unida a un modo de concebir la Ciencia que se ha dado a lo largo de la Historia y que, en cierta medida, fue el dominante. Se trata del que pretendía encontrar los elementos explicativos exclusivamente en los sistemas simples mecánicos y por lo tanto todos los fenómenos debían ser explicados en términos de esos sistemas.

A este respecto, podemos mencionar todas las teorías que usaron los fluidos imponderables o que despistaron y confundieron con ellos a lo largo de más de un siglo. La electricidad, el magnetismo, el calor, las reacciones químicas, el comportamiento de la luz como fluido radiante, son ejemplos de grupos de fenómenos que fueron explicados en términos de esos fluidos imponderables. Cuando se narra su Historia se suele decir que fueron superados por los sistemas mecánicos posteriores cuando se los interpretó por medio de sistemas más cercanos a la Mecánica racional. Pero, sin embargo, se suele omitir que ellos mismos fueron tratados como sistemas mecánicos muy simples. Si se desea una comprobación de esta afirmación sugiero que se acuda a los últimos volúmenes de la *Mecánica Celeste* de Laplace, los que estudian la forma de explicar *teóricamente* las leyes de la Física terrestre, no la celeste, que hasta entonces sólo se habían conocido desde un punto de vista experimental.[8]

En esa obra, todo se ve como el despliegue de corpúsculos y de fuerzas a partir del cual se reconstruyen las leyes del mundo físico, en un proceso que sería un sueño ideal para un lakatosiano. Así, a partir de sistemas mecánicos extraordinariamente simples se daría cuenta de toda la realidad física que estaría convertida en diferentes clases de corpúsculos sometidos a condiciones dinámicas perfectamente conocidas. El «maremagnum» de corpúsculos de calóricos, electricidad, magnetismo, en forma de atmósferas que rodean

7. Fraassen, B. van (1980), p. 13

8. En la edición de Laplace, P. S. (1829-1839), estos tratamientos terrestres se encuentran a partir del volumen tercero.

a los átomos de materia sometidos a la más amplia variedad de fuerzas no debe confundir al lector. Al final de toda esa hojarasca se establece la interacción entre corpúsculos materiales que pueden servir de origen para una reconstrucción matemática cuyo último estamento es la formulación de la ley física correspondiente. El itinerario puede ser muy proceloso pero el punto de partida y el punto final son muy claros. El primero es un sistema mecánico elemental y el segundo una ley física.

Si avanzamos unas cuantas décadas, se observará que Maxwell realiza la misma actividad con el Eletromagnetismo. Las ecuaciones del campo electromagnético, que siempre se ponen como ejemplo de una novedad dentro de la Física dominada por la Mecánica,[9] se obtienen por medio de hipotéticos sistemas mecánicos simples que le sirven como punto de partida para deducir matemáticamente las cuatro célebres ecuaciones que ofrecen la primera unificación eficaz de fenómenos electromagnéticos y ópticos.

La Mecánica fue una opción muy poderosa que dejó a su espalda una gran cantidad de problemas, aunque ese «dejar atrás» no impidió que fuera fecunda y que inspirara una buena parte de la Filosofía de la Ciencia del siglo siguiente. Si no se le puede reprochar a la Ciencia haber realizado un reduccionismo mecanicista semejante, tampoco se le debe negar a la Filosofía, en parte inspirada en ese contexto, un derecho a explorar los problemas metodológicos que nacieron al lado de la Mecánica.

3. El tratamiento de la complejidad

Frente a esa imagen de simplicidad metodológica, hoy se destaca la importancia de la complejidad como si hubiera sido un hallazgo o una invención de la Ciencia de nuestro tiempo. Y esto se piensa porque se han dado desarrollos en las Ciencias Formales y en las Ciencias de la Naturaleza que hacen gala, con justicia sin duda, del nombre de *Ciencias de la complejidad*. Pero sería un error grave suponer que la Ciencia que se dio desde la revolución científica del barroco no tuvo que enfrentarse a problemas derivados de la complejidad. Por eso, es necesario advertir que la actitud científica reduccionista a la que he aludido en el anterior epígrafe es una

9. El paso del electromagnetismo por la mecánica se puede ver especialmente en Maxwell, J. C. (1890), pp. 451-513, que correponden a un trabajo titulado «On Physical Lines of Force».

parte de la actividad científica, pero no el todo. No se puede negar la utilidad que han tenido los análisis formales de la Filosofía de la Ciencia tradicional, pero tampoco se debe reducir el papel de la Filosofía a tales análisis.

La clave la ofrece el mismo van Fraassen en el mismo libro mencionado, ampliando el campo de interés de las teorías científicas cuando plantea que la verdadera importancia de la teoría, para el científico activo, reside en que es un factor del diseño experimental. Esto es el reverso de la imagen elaborada por la Filosofía de la Ciencia tradicional. En esta imagen, todo se subordina al propósito de conocer la estructura del mundo... Sea cual sea la validez de esa imagen (y ciertamente tiene algo de verdad), contrasta fuertemente con la actividad que Kuhn llama «Ciencia normal» e incluso con mucho de lo que es revolucionario. Los científicos pretenden descubrir hechos acerca del mundo, acerca de las regularidades en la parte observable del mundo. Para ello se necesita experimentación, en lugar de razonamiento y reflexión. Pero esas regularidades son sumamente sutiles y complejas, de modo que el diseño experimental es sumamente difícil. De ahí la necesidad de construir teorías para guiar la investigación.[10]

A pesar de la longitud de la consideración, la he dado completa para evitar ser acusado de cercenar las opiniones ajenas en beneficio de los argumentos propios. Bas van Fraassen, coherente con todo el resto de su pensamiento, termina apostando por el papel reductivo de las teorías pero acepta que es el resultado de la excesiva complejidad de todo el entramado experimental.

Precisamente fue en la experimentación donde la primera Filosofía barroca, que luego daría lugar a la Ciencia moderna, encontró las primeras complejidades. Hoy se destaca que estudiar el movimiento del ondear de una bandera o el crecimiento de un ser pluricelular supone enfrentarse con un sistema complejo que requiere herramientas nuevas. Y esto es verdad, por supuesto. Pero también lo es que un telescopio en manos de un astrónomo barroco o una máquina de vacío en manos de un filósofo experimental de la Royal Society no eran también sino sistemas complejos. Cuando Hauksbee comunicó a Isaac Newton que había encontrado unas luminiscencias misteriosas en una máquina, tuvo que reconocer que no sabía absolutamente nada de lo que ocurría en el funcionamiento de semejante aparato.[11] Que ahora podamos explicarlo con

10. Fraassen, B. van (1980), p. 73.
11. Heilbron, J. L. (1979), pp. 229-249.

la Ciencia contemporánea, no nos autoriza a suponer que para aquellos filósofos naturales no se trataba de un sistema complejo. La Historia de la instrumentación, así como la de la experimentación, ha sido sistemáticamente apartada de la reflexión filosófica, la mayor parte de las veces con argumentos menos inteligentes que los expresados por van Fraassen.

Dejando a un lado los que tenían una función matemática o servían para realizar observaciones, habitualmente denominados *ópticos*, los instrumentos fueron construidos muchas veces con una finalidad específica: la de poner de manifiesto alguna propiedad que en general estaba mal definida. Éste fue el caso de los termómetros, los barómetros y, en general, todos los instrumentos relacionados con los fenómenos caloríficos. La temperatura y la presión eran dos magnitudes paradójicas que no se adecuaban a los requisitos de las magnitudes habituales de la Mecánica. No bastaba ver, observar y anotar el desplazamiento de una columna de mercurio, o un índice en un globo de cristal para dar una razón del comportamiento de la magnitud. Los debates ilustrados acerca de las escalas termométricas son una prueba elocuente de la dificultad que tenía explicar la aparentemente simple dilatación producida por un calentamiento. Los trabajos de Gay-Lussac y de Dalton pusieron de manifiesto que tales comportamientos no eran tan simples como parecían, y que sólo se podían abordar por medio del diseño de experimentos donde los gases aparecían como sistemas complejos.[12]

Lo mismo se puede decir de toda la Física experimental ilustrada, la que se llevaba a cabo como una suerte de Física recreativa, y que, sin embargo, fue la base del conocimiento de la electricidad y el magnetismo hasta finales del siglo XVIII. En ese caso los experimentos eran la teoría, como defendía Heilbron ya hace veinte años. De tal forma que si se dio progreso en esas Ciencias, éste se produjo en el contexto del diseño de los experimentos como procesos complejos que involucraban sistemas muchas veces mal definidos. Éste es el caso de los que llevó a cabo Franklin, para mostrar la naturaleza eléctrica de los rayos en los que se implicaba un sistema tan complejo como el desarrollo de una tormenta.[13]

Pero, generalmente, se trata todo este periodo barroco e ilustrado de una forma sumaria, ya que la actividad de Laplace en su *Mecánica Celeste* tendía a enmendar la complejidad por medio de una traducción, si no de una reducción, de todos los procesos apa-

12. Puede consultarse el capítulo segundo de Fox, R. (1971).
13. Heilbron, J. L. (1979), pp. 339 y ss.

rentemente complejos en procesos mecánicos. Por eso, cuando se elabora la Historia de la complejidad se suele considerar que los siglos XVII y XVIII son tersos periodos de mecanicismo militante.

La formulación de las grandes teorías físicas del siglo XIX, fundamentalmente las teorías relacionadas con el calor, la electricidad, el magnetismo y la óptica cambiaron bastante la perspectiva acerca de lo que podían ser los sistemas complejos. Conviene mencionar que el desarrollo del interés por los sistemas complejos también se dio en el contexto del nacimiento de las Ciencias sociales, aunque esto sería merecedor de otro extenso análisis y me voy a limitar a los problemas que emergieron en las Ciencias Físicas.

Si se toma en cuenta el trabajo de W. Weaver *Science and Complexity*, publicado en 1948, que siempre se ha considerado como un precedente en el tratamiento de la complejidad durante el último medio siglo, se podría establecer una clasificación sumaria de la Historia de la Física en tres periodos. Según la interpretación de Weaver, el primero abarca desde principios del XVII hasta finales del siglo XVIII, es decir, la época laplaciana. En ese tiempo, todo el desarrollo de la Ciencia estaría dominado por la vigencia de los sistemas simples, mecánicos, en los que todo se podría reducir a una causalidad entre elementos de carácter mecánico. Todos los problemas, tanto de carácter científico como de talante filosófico y metodológico, tendrían ese carácter mecánico en el que los individuos, o sucesos individuales, son perfectamente identificables en su relación con el tiempo. En el segundo periodo tendrían cabida los sistemas correspondientes a la *complejidad desorganizada* que, a su juicio, se definen a través de la Termodinámica clásica y la Mecánica estadística a partir de la segunda mitad del siglo XIX. Los científicos relacionados con este movimiento serían fundamentalmente Carnot, Clausius y Boltzmann. Según esto, los modelos de la Mecánica estadística se extendieron a otras Ciencias nacidas posteriormente como la biología y la Mecánica cuántica. Por último, el tercer periodo sería el actual, nacido a partir de la segunda guerra mundial, y correspondería al de la *complejidad organizada*. Los nuevos problemas que se abordan desde ella son demasiado complicados como para ser tratados por cualquier clase de Mecánica, sea racional o estadística. Según Weaver, son problemas biológicos, médicos, psicológicos, económicos y políticos, que no pueden ser tratados con modelos lineales y causalistas y requieren nuevas estrategias metodológicas.

Dejando a un lado las buenas intenciones de Weaver, su capacidad para anticiparse a problemas que se plantearon de una forma más matizada a partir de los años setenta, y su intuición para ver

las preocupaciones éticas y políticas sobre la Ciencia, me voy a detener en su clasificación de estos tres periodos que supuestamente ha atravesado la constitución de la complejidad.

Por lo que he adelantado, durante el primer periodo al que hace referencia lo que se dio fue una cierta propensión a reducir cualquier sistema a elementos simples. Los sistemas complejos quedaron relegados al ámbito de la instrumentación y la experimentación, mundo confuso en el que pocas veces entraron los filósofos de ése y otros muchos periodos. El sistema complejo de un experimento más tarde sería visto como el prototipo de «caja negra» del que cabe sólo extraer información acerca de los datos que supuestamente son necesarios para la elaboración de las teorías. Pero un termómetro, un barómetro o una pila voltaica no ofrecían datos, ni medidas que pudieran ser explicados por medio de una teoría reductora y, sin embargo, no podemos pensar que no tuvieran una función en el desarrollo de la Ciencia del momento.

Pero aparte de estas cuestiones, conviene fijarse en el segundo periodo que —a mi juicio— es fundamental para entender el papel de la explicación científica en la Ciencia contemporánea más allá de las simplificaciones mecanicistas ilustradas. En ese periodo no sólo intervinieron los actores que menciona Weaver, cuyos trabajos científicos sitúa en la mitad del siglo XIX (las imprecisiones cronológicas son poco importantes ya que él mismo no era un historiador profesional). En realidad, los problemas comenzaron una generación antes, en plena generación de la confianza en la Mecánica, en la generación laplaciana.

En el primer cuarto de siglo XIX, los problemas planteados por las Ciencias Físicas, ya de una cierta complejidad como es el caso de la acción del calor sobre los gases o la transmisión del calor en los diferentes tipos de sólidos, se intentaron resolver de dos modos fundamentales. O bien suponiendo que la acción del calor se podía *representar* por medio de sistemas mecánicos muy sencillos pero muy numerosos —como era el caso de la representación cinético corpuscular de los gases— o bien que era necesario renunciar a tal representación y sólo había que expresar los fenómenos implicados en el proceso por medio de unas ecuaciones que no presupusieran ninguna entidad subyacente. La Ciencia debía ser una actividad para *describir* los fenómenos sin ir más allá de ellos.

La primera posición fue adoptada por muchos laplacianos, como Gay-Lussac, Petit, Dulong y otros, mientras que el representante más reconocido del segundo punto de vista es Fourier. Muchas veces se presenta la polémica entre los dos puntos de vista

como el enfrentamiento entre un cierto tipo de realismo frente a un positivismo emergente que influyó de forma decisiva en la vida científica del siglo. Pero además de esa vertiente, la polémica ofrece otras igualmente interesantes. Aquí me gustaría mencionar la que opone la *descripción* a la *explicación*.

Efectivamente, una gran parte de los físicos continentales del siglo XIX, primordialmente franceses y centroeuropeos mantuvieron que la finalidad de la Ciencia era describir fenómenos y no ofrecer explicaciones de ellos. Suponían que *explicar* era una actividad peligrosamente anclada en hipótesis atomísticas. Por el contrario, *describir* permitía liberarse de esas hipótesis dado que el formalismo matemático se limitaba a traducir los fenómenos.

4. En defensa de la explicación en el mundo de la complejidad

El momento más relevante de la polémica entre los partidarios de la Ciencia como explicación y los defensores de su reducción a una descripción se dio en la segunda mitad del siglo XVIII. El representante más relevante de la segunda opción fue Gustav Kirchhoff, conocido por haber difundido en Centroeuropa la idea de que el único objeto de la Ciencia debe ser la descripción más sencilla y completa de los fenómenos. A pesar de la amplia aceptación que esta prescripción metodológica tenía en la comunidad científica germanoparlante, no todos la aceptaban. Uno de sus miembros más amantes de la polémica, el prestigioso físico austriaco Ludwig Boltzmann, no la compartía. Por el contrario defendía que el propósito de la Ciencia era construir explicaciones, consistentes en la reducción a los conceptos más simples y basados en las ideas científicas más comúnmente aceptadas. Boltzmann estaba fuertemente influido por la Metodología británica, especialmente la maxwelliana que acudía a modelos y analogías como guía eficaz para hacer nuevos descubrimientos. Pero con su actitud no estaba defendiendo el mundo de la simplicidad matemática ilustrada, sino que pretendía explicar las leyes que más habían incomodado a los físicos de su tiempo. Me refiero a las leyes sobre la irreversibilidad de los procesos naturales que postulaban una cierta dirección en la evolución de los sistemas, incomprensible desde el punto de vista de la Mecánica racional.

Boltzmann para eso necesitaba que «la tarea de la Ciencia [fuese] aclarar las cosas más complicadas a partir de las más simples; o bien, si se quiere representar lo complejo por medio de imágenes

claras obtenidas en el ámbito de los fenómenos simples».[14] La Ciencia únicamente puede disolver lo complejo en partes más simples pero idénticas y reducir lo que es complicado a leyes más fundamentales. Sin embargo, los elementos más simples «a los que todo se reduce permanecen inexplicables para siempre». En este aspecto se alineaba con la más pura tradición metodológica que ha defendido la prioridad de la explicación basada en verdades primitivas.

En realidad, la referencia fundamental de toda explicación volvía a ser la Mecánica, que ocupaba un lugar central en la Filosofía de la Ciencia de Boltzmann. Así lo expresaba en su obra *Vorlesungen über die Prinzipie der Mechanik*: «Seguramente estaría todo mucho más claro si pudiéramos explicar no sólo todos los fenómenos de movimiento en los cuerpos sólidos, líquidos y gaseosos, sino también el calor, la electricidad, el magnetismo y la gravitación por medio del concepto de movimiento de puntos materiales en el espacio; lo que quiere decir, más por medio de un principio simple y uniforme, que a través de una colección de conceptos completamente heterogéneos como la temperatura, la carga eléctrica, el potencial, etc., para cada una de estas materias».[15] Con estas reflexiones parecía regresar al mundo de la vieja Mecánica, aunque realmente no fuera propiamente un regreso, sino un paso adelante que destruiría la vieja confianza de la Mecánica en la reversibilidad de los procesos naturales. Por eso, Boltzmann siempre preconizó que estaba fundando una nueva Mecánica para entender los procesos naturales relacionados con el calor.

En la reunión de científicos alemanes celebrada en Munich en 1899, Boltzmann ya se presentaba a sí mismo como «alguien aparte, quien, al contrario de los modernos (es decir, los partidarios de la Ciencia como *descripción*), veneraba a los antiguos clásicos», y subrayaba los enormes cambios que había sufrido la Ciencia desde sus tiempos de estudiante. Pero «los antiguos», comentaría más tarde Lorentz, cuando en su conferencia en memoria de Boltzmann revisaba el impacto que su Física había tenido en los físicos modernos, «retornan a una vida nueva y vigorosa».

En todo caso, el logro más importante de Boltzmann no fue mantener la vigencia de la Mecánica, entendida como Ciencia básica para interpretar la Naturaleza, sino reducir «la incomprensible irreversibilidad» (dictaminada desde S. Carnot como algo intrínseco a las transformaciones calor/trabajo, es decir, a cualquier gene-

14. Boltzmann, L. (1986), p. 193.
15. Boltzmann, L. (1897), p. 137.

ración de trabajo) a conceptos conocidos, a un teorema de teoría de la probabilidad. Esto implicaba una ruptura con el tradicional principio de la descripción mecánico-determinista que tan frecuentemente se había encargado de enfatizar. De forma paralela a Lagrange, quien consiguió reducir la Mecánica, la Mecánica racional, a un problema de la matemática, Boltzmann recondujo la irreversibilidad a un problema de la teoría de probabilidades.

Con tal reducción provocó la ruptura con la vieja tradición determinista y provocó una profunda transformación en la Ciencia. En definitiva, colaboró decididamente para abandonar lo que se ha dado en llamar la imagen clásica del mundo que ofreció la Física durante dos siglos. La importancia de la aproximación probabilística de Boltzmann se puede comparar con la que en su tiempo tuvo la teoría gravitatoria de Newton.

El físico austriaco provocó el paso de una concepción dinámica de la Naturaleza a una estadística, que no suponía solamente un cambio de método, sino también de nuestros puntos de vista acerca de la estructura conceptual de las leyes físicas. Este cambio también fue señalado por Hasenöhrl, discípulo de Boltzmann, cuando se preguntó: «¿qué es el azar? ¿Es el azar sólo necesidad oculta, la ausencia de una razón suficiente, la excepción de la ley de la causalidad como ha sugerido recientemente un conocido colega?».[16] El joven Erwin Schrödinger más tarde, en sus memorias, rememoraría el discurso inaugural de Hasenöhrl «pronunciado sin solemnidad en el otoño de 1907 en el antiguo salón de actos del viejo Instituto de Türkesstrase, explicándonos de manera resumida, con palabras sencillas y claras pero entusiastas, las ideas básicas de los trabajos que desarrolló Boltzmann a lo largo de su vida. La descripción dejó en mí una profunda impresión intelectual que ya nunca abandonó mis pensamientos, y nada en la Física me pareció más importante que las ideas de Boltzmann a pesar de Planck o Einstein».

5. La justificación de un trabajo científico

Aquí, sin embargo, no me interesa destacar sus triunfos posteriores sino el juego que tuvo a la hora de ampliar el horizonte de qué significaba la explicación en la Ciencia. Más que destacar las reacciones personales de los discípulos o seguidores de Boltzmann, me parece relevante resaltar el origen de sus ideas. El cómo y por

16. Hasenöhrl, F. (1874-1915), pp. 14 y ss.

qué acudió a las ideas probabilísticas para dar una versión de las leyes de la irreversibilidad termodinámica.

A finales del siglo XIX, la Física alemana como ya he dicho estuvo dominada por los partidarios de la Ciencia como una descripción. Pero a pesar de ese dominio, no dejaron de existir representantes de tendencias más mecanicistas. Incluso hubo influencias de la Física británica a través de traducciones llevadas a cabo por el ingeniero August Krönig.[17] Tales traducciones llegaron a manos de Rudolp Clausius quien en 1857 ya dio una primera interpretación cinético corpuscular (y por lo tanto mecanicista) del calor. Otros físicos, bajo la dirección de Andreas von Ettinghausen en el Instituto de Física de Viena, se preocuparon activamente de problemas relacionados con la Física molecular. Todos ellos pertenecían a la escuela de Viena y en su contexto se tradujo, estudió y difundió la obra electromagnética del escocés Maxwell. Posteriormente, Josef Stefan sucesor de Ettinghausen en la dirección del Instituto, continuó tanto los trabajos científicos como su interés por la Física británica. En ese contexto se formó como científico y trabajó Ludwig Boltzmann a partir de 1865.

El interés por la Física británica no se refería exclusivamente al Electromagnetismo. También atendía a los desarrollos cinético-corpusculares y a las aplicaciones de la Estadística al tratamiento de los gases. Para los que eran partidarios de la Ciencia como *descripción*, la formulación del segundo principio de la Termodinámica estaba bien construida por medio de las ecuaciones de Clausius. Pero había unos pocos que deseaban una *explicación* de la misma a través de elementos mecánicos sencillos. La mayor parte de esos pocos estaba en Viena y uno de ellos era Boltzmann que se planteó la cuestión como un problema prioritario.

Para lograr la interpretación mecánica de la segunda ley se exploraron varios caminos. Científicos tan prestigiosos como Helmholtz y Clausius escribieron trabajos muy meritorios e interesantes en los que se deseaba reducir el sistema irreversible del que se tratase, y por lo tanto complejo, a sistemas simples. Boltzmann exploró otras posibilidades explicativas. En un principio, hasta 1868, trabajó bajo las sugerencias de Thomson y Tait y trató de aplicar el principio de mínima acción de Maupertuis. Todavía años más tarde consideraba que era un camino que podía dar sus frutos aunque no los que él deseaba. Cuando el joven físico japonés Hantaro Nagoaka visitó Munich en 1884, tuvo la oportunidad de escuchar las

17. Brush, S. (1976), p. 165.

conferencias de Boltzmann. «Hablaba sobre el sistema monocíclico y el principio de Maupertuis. Todas las leyes naturales, como por ejemplo la ecuación de Lagrange, el principio de conservación de la energía, la segunda ley de la Termodinámica, la dispersión de la luz y la ley de Newton, pueden derivarse del principio de mínima acción, que se puede considerar como el principio más general de la Física.»[18]

Posteriormente, e influido por Helmholtz, trabajó en la aplicación de un principio más abstracto todavía, como el principio de Hamilton. En 1872, Boltzmann informó por primera vez acerca de los avances que había realizado en la resolución del problema de la irreversibilidad. Para asegurarse la prioridad envió una breve comunicación al *Wiener Anzeiger* que se imprimió el 8 de febrero.[19]

Fue la ocasión elegida por Boltzmann para presentar, junto a sus nuevos y abundantes resultados, también su famoso teorema-H, que podía estar relacionado con la entropía. Su formulación contenía la distribución de velocidades de Maxwell, así como una nueva definición de entropía también para condiciones de no-equilibrio, lo cual no era posible en las aproximaciones fenomenológicas y meramente descriptivas del problema. Pero para la mayoría de los científicos pasó desapercibido el aspecto fundamental de la teoría: que el comportamiento irreversible de la función-H únicamente se introducía en la formalización a través de la condición implícita del desorden molecular. Es decir, pasaba desapercibida la primera condición para que se diera una complejidad real, es decir, la desaparición de los comportamientos individuales.

En sus trabajos posteriores, Boltzmann sumó resultados científicos cada vez más interesantes. En 1877, tras haber explorado todas las posibilidades de un reduccionismo mecánico estricto fue capaz de deducir la entropía de gases ideales monoatómicos con la ayuda de valores medios en un espacio de fases. En ese mismo año ofreció su concepción acerca de la entropía en toda su generalidad como una medida de la probabilidad de los estados macroscópicos, y, en consecuencia, como una nueva clase de legalidad física opuesta a la visión tradicional de leyes absolutas.

La Estadística de Boltzmann fue aceptada después de su defensa, ante las objeciones que en las décadas siguientes le hicie-

18. Citado en Meyenn, K. von (1993), p. 148.

19. El artículo de L. Boltzmann de 1872 fue publicado finalmente en *Sitzungsberichte der Akademie der Wissenschaften*, Wien, II, v. 66, p. 275, con el título «Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen».

ron fundamentalmente Poincaré y Zermelo. Finalmente, su oponente más acérrimo, Max Planck, también pareció convencerse del valor de las ideas del físico austriaco, especialmente cuando pudo apreciar las ventajas del método de Boltzmann en la aplicación del segundo teorema de la Termodinámica a los fenómenos de radiación. La subdivisión de Boltzmann del espacio de fases en unidades discretas, usada como un artificio técnico para contar las complexiones correspondientes a un estado macroscópico, se convirtió en una estrategia fundamental en la teoría de Planck cuando éste trató de derivar la ley empírica de la radiación de los primeros principios.

Los procedimientos de Boltzmann también sustituyeron a los elementos esenciales que Planck venía utilizando en sus razonamientos formales. Como contrapartida, el gran éxito de la Teoría cuántica de Planck otorgó a la Estadística de Boltzmann un lugar central en el cuerpo de la Física que, todavía hoy, mantiene. Otro de los resultados que trajeron consigo las disputas sobre la fundamentación estadística de la Termodinámica, fue el reconocimiento de que, de igual forma que la probabilidad, la hipótesis del desorden molecular —o el concepto correspondiente de «radiación natural» en la teoría de la radiación— era necesaria para dar una explicación completa de la irreversibilidad. Esta suposición, así como el supuesto de una estructura discreta del espacio de fases eran incompatibles con los principios de la Mecánica clásica, y situaron en primer término el problema de la importancia de los sistemas complejos.[20]

6. **Proyección ulterior**

No es necesario llegar a nuestro final de siglo, como sugiere Weaver, para comprobar la existencia de un interés científico por los sistemas complejos. Ya la Ciencia barroca e ilustrada tuvo que entenderse con ellos, aunque siempre fueran tratados de forma oblicua dada su intraducibilidad en términos mecánicos simples. Además, creo que la exposición anterior ha ilustrado cómo una de las polémicas centrales de la Ciencia del siglo XIX se dio en torno a un problema de complejidad donde la estructura secuencial que se ha analizado anteriormente queda disuelta, es decir, aquella estructura donde se identifican perfectamente los individuos que intervienen

20. Meyenn, K. von (1993), p. 151.

en la interacción y donde se considera real la transmisión de una determinada propiedad. A partir de entonces, las condiciones de contorno del problema variaron y las propiedades estudiadas hablaban de sistemas y no de individuos.

Pero en particular, deseo señalar algo más que no se refiere al nivel de la actuación de los científicos a la hora de construir el conocimiento, más bien tendría interés desde un punto de vista filosófico y metodológico: Así como el reduccionismo antiguo mecanicista tuvo su origen en el interés por convertir las Ciencias en lenguajes explicativos, la emergencia de los sistemas complejos, como los termodinámicos y los estadísticos del siglo XIX, también estuvo determinada por un ideal explicativo. La interpretación de la Ciencia como descripción pudo parecer filosóficamente correcta durante un tiempo, pero resultó ser una actitud profundamente conservadora. La interpretación de la Ciencia como una actividad explicativa, con todos los riesgos que conllevaba, resultó ser innovadora, capaz de una mayor potencia heurística y filosóficamente más interesante.

Referencias bibliográficas

BLAY, M. (1992), *La naissance de la mécanique analytique. La science du mouvement au tournant des XVIIe et XVIIIe*, P.U.F., París. Incluye un prefacio de Jacques Merlau-Ponty.

BOLTZMANN, L. (1872), «Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen», *Sitzungsberichte der Akademie der Wissenschaften*, Wien, II, v. 66, pp. 275-370.

BOLTZMANN, L. (1897), *Vorlesungen über die Principe der Mechanik*, Barth, Leipzig.

BOLTZMANN, L. (1986), *Escritos de Mecánica y Termodinámica*, edición de Javier Ordóñez, Alianza Editorial, Madrid.

BRUSH, S. (1976), *The Kind of Motion we Call Heat*, North-Holland, Amsterdam.

FOX, R. (1971), *The Caloric Theory of Gases from Lavoisier to Regamult*, Clarendon Press, Oxford.

FRAASSEN, B. VAN (1980), *The Scientific Image*, Clarendon Press, Oxford. Vers. cast.: *La imagen científica*, Universidad Nacional Autónoma de México, México, 1996.

HASENÖHRL, F. (1913), «Festrede». En: *Bericht über die Errichtung eines Denkmals für Ludwig Boltzmann im Arkadenhof der Wiener Universität*, editado por Dr. F. Hasenöhrl, Viena, pp. 7-20.

HEILBRON, J. L. (1979), *Electricity in the 17th and 18th Centuries. A Study of Early Modern Physics*, University of California Press, Berkeley. [Hay una edición en rústica de 1999].

KOYRÉ, A. (1994), *Pensar la Ciencia*, Paidós-I.C.E./U.A.B., Barcelona. Introducción de C. Solís, pp. 9-43.

LAKATOS, I. (1978), *The Methodology of Scientific Research Programmes*, edición de J. Worrall y G. Currie, Cambridge University Press, Cambridge. Vers. cast. de Juan Carlos Zapatero: *La Metodología de los programas de investigación científica*, Alianza Editorial, Madrid, 1983.

LAPLACE, P. S. (1829-1839), *Traité de mécanique celeste*, Bachelier, París (edición en 4 volúmenes).

MAXWELL, J. C. (1890), *The Scientific Papers of J. C. Maxwell*, edición de W. D. Niven, Cambridge University Press, Cambridge.

MEYENN, K. VON (1993), «Dynamical and statistical conceptions in Boltzmann's Physics», en BATTIMELLI, G., IANNIELLO, M.ª G. y KRESTEN, O. (eds), *Proceedings of the International Symposium on Ludwig Boltzmann, (Rome, February 9-11, 1989)*, Verlag der Österreichischen Akademie der Wissenschaften, Viena, pp. 141-162.

TRUESDELL, C. (1968), *Essays in the History of Mechanics*, Springer Verlag, Berlín. Vers. cast. de Juan Carlos Navascués Howard y Enrique Tierno Pérez-Relaño: *Ensayos de Historia de la Mecánica*, Tecnos, Madrid, 1975.

WEAVER, W. (1948), «Science and Complexity», *American Scientist*, v. 36, p. 536.

CAPÍTULO 3

EXPLICACIÓN, UNIFICACIÓN Y SUBSUNCIÓN[1]

JOSÉ A. DÍEZ CALZADA

En diversos lugares, W. C. Salmon, uno de los principales defensores del modelo causal de la explicación, defiende la compatibilidad de este modelo con el modelo unificacionista:

> «Estos dos modos de ver la explicación *no son incompatibles* entre sí; cada uno ofrece un modo razonable de analizar la explicación. De hecho, pueden ser tomados como representando dos aspectos diferentes pero compatibles de la explicación científica» (1989, p. 183).
>
> «[Estos dos enfoques] se han desarrollado hasta el punto en que pueden coexistir pacíficamente como dos aspectos distintos de la explicación científica» (1992, p. 39).
>
> «No cuestiono la posibilidad de una teoría satisfactoria de esta clase [unificacionista]; pienso que podría *complementar* en lugar de *entrar en conflicto* con la explicación causal» (2002*a*, p. 105).

El objeto de esta contribución es explorar las diversas posibilidades de esta tesis compatibilista, determinar cuál es el sentido más plausible y presentar un modelo que nos parece especialmente adecuado para expresar dicho sentido.

1. Explicación causal y explicación como unificación

La idea central del modelo causal de explicación es que «explicar un acontecimiento es proporcionar información acerca de su historia causal» (Lewis, 1986, p. 217). La explicación, sin embargo,

1. Este trabajo forma parte de los proyectos de investigación PB98-0495-C08-07 y PFF2000-1073-C04-04, DGES, Ministerio de Educación. Agradezco a Wesley C. Salmon, Merilee H. Salmon, Javier Echeverría, Wenceslao J. González, Javier Ordóñez, Joan Pagés y Manuel Pérez sus comentarios y críticas a versiones previas de este trabajo.

no debe informar acerca de la historia causal *completa*. La historia causal completa de un acaecimiento incluye multitud de otros acaecimientos, y la mención a todos ellos no sólo no es necesaria sino que sería incompatible con una buena explicación. Quienes, como Salmon, reivindican el carácter objetivo de la explicación causal, aceptan que el contexto determine qué parte de la historia causal es la explicativamente relevante. La explicación causal es objetiva porque, si se tiene éxito en la reconstrucción de una noción objetiva de *causalidad* (como pretende, p. ej., Salmon con su análisis de *estructura causal completa*), los antecedentes causales que el contexto pragmático determina como explicativamente relevantes en dicho contexto son parte de una estructura causal objetiva. Así, el carácter óntico u objetivo no se pierde por reconocer a la vez la naturaleza intrínsecamente epistémica y pragmática de la explicación.

El modelo causal resuelve satisfactoriamente los principales contraejemplos a que se enfrentaba el modelo inferencial de Hempel. Simetría: la altura del mástil explica la longitud de la sombra, pero no a la inversa, puesto que la primera pertenece a la historia causal de la segunda y no a la inversa. Irrelevancia: la ingestión de pastillas anticonceptivas por parte de Juan no explica su no embarazo pues no pertenece a la historia causal de su no embarazo. «Horquillas» o antecedente común: el descenso brusco del barómetro no explica la tormenta sino que la bajada brusca de presión explica ambos puesto que ambos son causados por ella. Según algunos críticos, este modelo tiene problemas para dar cuenta de la explicación de hechos particulares probabilistas, de la explicación de hechos generales y de los casos de supuestas explicaciones no causales. Diferentes defensores del modelo afrontan cada uno de estos casos de modo específico, pero siguiendo en general una de dos estrategias: negando que se trate de genuina explicación, o ampliando no trivialmente la noción de explicación causal para dar cuenta del caso. El reto principal, sin embargo, sigue siendo proporcionar un análisis satisfactorio de la noción de causalidad (siempre que no se tome como noción primitiva). Algunos lo hacen en términos de otras nociones modales (p. ej., Lewis, 1973, utiliza la noción de dependencia contrafáctica). Otros filósofos de orientación más empirista que consideran las nociones modales primitivas sospechosas, intentan otras estrategias en términos no modales (p. ej., Salmon, 1981, 1984, 1994 y 1997, y Dowe, 1992 y 1995). Para estos últimos es especialmente importante que su análisis dé a la vez respuesta adecuada al problema de Hume.

La intuición básica que inspira el modelo de explicación como unificación es que explicar consiste en «reducir la cantidad de fenómenos independientes que tenemos que aceptar como últimos» (Friedman, 1974, p. 18). Aumentamos nuestra comprensión del mundo reduciendo el número de supuestos básicos que dan cuenta de los fenómenos. Las leyes de Newton explican las de Kepler porque, además de implicarlas, reducen la cantidad de regularidades que se aceptan independientemente unas de otras: antes de la explicación, las leyes de Kepler y, por ejemplo, la de Galileo eran aceptadas independientemente unas de otras, después no; la reducción de la temperatura a la energía cinética media amplía todavía más ese proceso de unificación explicativa. Contrariamente, la mera conjunción de, por ejemplo, las leyes de Kepler con la de Boyle no es una explicación de las primeras porque no produce ese efecto unificador, no permite simplificar la cantidad de supuestos primitivos. Esta noción de explicación está esencialmente relativizada a un cuerpo K de creencias aceptadas en un momento dado, y exige una elucidación precisa de la *independiente aceptabilidad* entre creencias. Quien ha desarrollado con más detalle este modelo ha sido Ph. Kitcher mediante sus nociones de *patrón argumentativo* y *«almacén» explicativo* (*explanatory store*) (1981, 1989, 1993), caracterizando el poder unificador en función de: *a*) la cantidad de fenómenos derivados por el conjunto de patrones argumentativos, *b*) el rigor de los patrones, y *c*) el número de patrones.

El modelo unificacionista tiene su propio modo de dar cuenta de los casos de simetría, irrelevancia y horquillas. La estrategia es mostrar que, dadas dos inferencias alternativas, será explicativa la que pertenezca a la sistematización más unificadora, y que esta comparación arroja en los casos en consideración los resultados intuitivamente esperados. Horquillas: la sistematización que deriva la bajada del barómetro y la ocurrencia de la tormenta a partir del descenso de la presión es más unificadora que la que deriva la ocurrencia de la tormenta a partir del descenso del barómetro. Simetría: la altura del mástil explica la longitud de la sombra y no al revés pues dadas dos sistematizaciones, una que contenga inferencias que parten de la altura del mástil y conducen hasta la longitud de la sombra, y otra que contenga inferencias que proceden al revés, la primera es más unificadora que la segunda; si la segunda no tiene otro tipo de inferencias, pierde algunas conclusiones pues no podrá establecer, p. ej., la altura de mástiles de noche, o en días nubosos, etc.; para recuperar esas conclusiones debería introducir nuevos patrones argumentativos. Irrelevancia: una sistematización que contiene derivaciones del no embarazo de Juan usando como premisa que Juan toma pastillas

anticonceptivas no puede ser la mejor pues, o bien no explica que otros varones que no toman pastillas tampoco se quedan embarazados, o si lo explica deberá introducir nuevos patrones que también se aplicarán a Juan, con lo que podríamos prescindir de las primeras inferencias obteniendo una sistematización con menos patrones, más unificada. Como reconoce Kitcher (1989, p. 483), estos dos últimos casos requieren que las sistematizaciones no usen predicados del tipo de «verdul», esto es, predicados no proyectables.

Aunque el modelo unificacionista se aplica más naturalmente a explicaciones de hechos generales, se puede aplicar también a hechos particulares considerando estos casos derivativos de aquéllos: «La pregunta '¿por qué este objeto particular se comporta de este particular modo?' es transformada en la pregunta '¿por qué objetos ideales de este tipo general exhiben esas propiedades?'» (Kitcher, 1989, p. 453). Por otro lado, para este modelo no representan en principio ningún problema las explicaciones de hechos probabilistas, siempre que se puedan inferir de ciertos patrones, ni las explicaciones intuitivamente no causales, pues no se compromete con ningún tipo de mecanismo específico que expresen los patrones argumentativos. El modelo unificacionista debe afrontar, principalmente, dos retos. El primero es el de proporcionar un criterio adecuado de ponderación entre los parámetros (a)-(c) que determinan la simplicidad comparativa. El segundo es el de recoger un sentido suficientemente fuerte de la objetividad de las explicaciones. Acabamos de ver que para descartar, por ejemplo, la explicación del no embarazo de Juan a partir de su ingestión de anticonceptivos, debemos excluir que la sistematización use predicados no proyectables.[2] Sin embargo, la elucidación independiente de cuándo un predicado es *proyectable* o *natural* constituye un problema en sí mismo. Para unos ello nos retrotrae al problema de la causalidad, en cuyo caso estaríamos implícitamente ante análisis causalistas, que serán realistas o antirrealistas dependiendo de cómo se analice a su vez la causalidad. Para otros, tal elucidación sólo puede ser epistémica o pragmática, en este caso el análisis suele ser, aunque no siempre se reconozca, de corte antirrealista.[3]

2. Es fácil ver que si usamos predicados como «tomar anticonceptivos, o no tomarlos y ser varón», o «estar embrujado, o no estarlo y tener estructura NaCl» (para el caso de la explicación de la disolución de terrones de azúcar a partir de haber sido embrujados), se pueden simplificar las sistematizaciones que generan esas supuestas explicaciones espurias.

3. Kitcher, por ejemplo, afirma que los predicados deben ser proyectables «desde la perspectiva de *K*» (1989, p. 483). Aunque él rechazaría el cargo de antirrealismo, más adelante veremos que el modo como él reconstruye la causalidad y las dependencias objetivas no es todo lo robusto que pretenden los realistas.

Vamos ahora a explorar las diversas posibilidades de compatibilidad entre ambos modelos. Para ello utilizaremos ocasionalmente el esquema que van Fraassen (1977) ofrece de los contextos explicativos, esquema que pone de manifiesto los diversos parámetros involucrados en una explicación y que cada modelo fija de un modo específico. Cada uno de tales contextos consiste al menos en una *pregunta por qué* (*why question*), por ejemplo, «¿por qué Adán comió la manzana?», «¿por qué el alcalde contrajo paresis?». Esta pregunta es el *tema*. Con el tema solo, sin embargo, no basta, pues no expresa el aspecto por el que se demanda explicación. En el primer caso, por ejemplo, tenemos tres alternativas: «¿por qué *Adán* (y no Eva, o la serpiente, o...) comió la manzana?», «¿por qué Adán *comió* (y no rechazó, o ...) la manzana?», «¿por qué Adán comió *la manzana* (y no una pera, u otra manzana, o ...)?». Así, junto al tema debe incluirse (usualmente de modo implícito determinado por el contexto), una clase de alternativas, la *clase de contraste*, frente a las que se contrapone el hecho por cuya razón se inquiere: «¿por qué a, contrariamente a β_1, β_2, ...?». En opinión de van Fraassen, el tema a y la clase de contraste X no bastan, sin embargo, para identificar completamente la demanda de explicación pues, incluso fijada X, puede haber varios tipos de respuesta adecuada dependiendo de qué relación se considere en ese contexto que es la relevante para poder considerar a una respuesta una explicación adecuada.[4] Hasta que el contexto no determina *el tipo* de respuesta que se considera explicativa la demanda de explicación está indeterminada. Junto con el tema y la clase de contraste hay que incluir entonces una *relación de relevancia explicativa*, R.[5] Así, podemos representar un contexto explicativo mediante una terna $Q=<a, X, R>$, donde a es el tema y X y R son, respectivamente, la

4. Por ejemplo, supongamos que pedimos razón de que Juan viniera a la fiesta contrariamente a que se quedara en casa o se fuese al cine. Una respuesta puede ser que Juan salió de su casa un rato antes, caminó en cierta dirección por cierta ruta, se detuvo en un portal, subió a un piso y el piso es donde tiene lugar la fiesta. Pero en muchos contextos no se aceptaría esta respuesta como explicación porque se pedían razones *de otro tipo*. La respuesta no es del tipo que el contexto considera relevante, mientras que sí lo sería, por ejemplo, que Juan quería ver a Luisa y sabía que Luisa iba a ir a la fiesta.

5. R sólo determina *el tipo* de respuesta considerada relevante, no la respuesta misma, pues hay varias respuestas que pueden ser relevantes en el contexto; p. ej., en el caso mencionado podría quizás considerarse también relevante como respuesta el que Juan quiere ir al menos a una fiesta cada mes, que ese día es fin de mes y que ese mes todavía no ha ido a ninguna. Por otro lado, van Fraassen no impone ninguna restricción a R, pudiendo cualquier respuesta (que tenga que ver con el tema) ser una explicación *adecuada dado un contexto apropiado*. Kitcher y Salmon (1987) muestran que ello es inaceptablemente liberal.

clase de contraste y la relación de relevancia explicativa determinadas por el contexto. En algunos contextos, como los científicos en periodos de ciencia normal, están fijados con bastante rigidez, pero en otros pueden ser muy variables.

2. Posibilidades de compatibilidad

Las principales posibilidades de compatibilidad de dos conceptos son, por orden de fuerza, las siguientes: dependencia conceptual fuerte o total de uno respecto de otro, esto es, identidad conceptual; dependencia conceptual débil, esto es, inclusión conceptual; solapamiento conceptual, esto es, núcleo conceptual común sin dependencia mutua; diversos tipos de solapamiento extensional (coextensividad, inclusión extensional o solapamiento sin inclusión); complementariedad extensional sin solapamiento. Veamos ahora estas posibilidades para el caso de los conceptos «explicación causal» y «explicación como unificación», en adelante, EC y EU respectivamente.

1) EC se reduce conceptualmente a EU. La afirmación de van Fraassen de que podemos definir la red causal como «cualquier estructura de relaciones que la ciencia describe» (1980, p. 124) apunta en esta dirección. Ésta parece ser también la posición defendida por Kitcher en algunos pasajes como éste: «el asunto crucial es que el 'debido a' (*because*) de la causalidad se deriva siempre del 'debido a' (*because*) de la explicación» (1989, p. 477, aunque una interpretación más débil podría limitarse a dependencia parcial). Quien defiende esta alternativa, sostiene que (independientemente de lo que parece preanalíticamente) no es posible proporcionar un concepto *satisfactorio* de causalidad independiente del de EU: «Recomiendo rechazar la idea de que hay verdades causales que son independientes de nuestra búsqueda de orden en los fenómenos» (*ibidem*, p. 497); «carece de sentido la noción de relevancia causal independiente de [la noción] de relevancia explicativa y no tiene sentido para la noción de relevancia explicativa, excepto que figure en la sistematización de la creencia en el límite de la investigación científica, en cuanto orientada por la búsqueda de la unificación» (*ibidem*, p. 499). Esta doctrina tiene ecos antirrealistas, pues construye un concepto metafísico a partir de otro epistémico, y resulta por tanto sospechosa a todo aquel que persiga un concepto más robusto de causalidad. Para algunos, además, supone una trivializa-

ción del concepto de causa ampliándolo inaceptablemente (Ruben, 1992, p. 211). Por otro lado, en la versión de Kitcher tiene la dificultad de que, como vimos, su concepto de EU presupone la distinción entre predicados proyectables y no proyectables, distinción de la que habría que dar cuenta en términos no causales (so pena de caer en circularidad). Por último, este sentido fuerte que estamos considerando es conceptualmente incompatible con la existencia de explicaciones no causales, cuya posibilidad era una de las motivaciones de los unificacionistas. Quien simpatice con la prioridad conceptual de EU debe optar, parece, por una interpretación más débil.

2) EC no se identifica con EU pero lo presupone conceptualmente: toda EC es, conceptualmente, EU (aunque la inversa no es cierta). Esto es, las EC son, conceptualmente, EU *de cierto tipo específico*: «*x* C-explica *y* syss *x* U-explica *y* y además ...». Esta alternativa puede tener al menos dos versiones, según el concepto de causa se derive del de sistematización unificadora o no. En el primer caso, la razón de que EC presuponga conceptualmente EU es que el concepto de causa se analiza en términos de EU. Ésta sería la versión correspondiente a la lectura débil de los anteriores pasajes de Kitcher y comparte algunas de las dificultades de la versión fuerte. No trivializa, en principio, el concepto de causa como aquella, pero a cambio debe especificar qué es lo que caracteriza el tipo concreto de unificación que «genera» relaciones causales, y lo debe hacer en términos no causales so pena de circularidad. En el segundo caso, se dispone de un concepto de causa independiente y el vínculo conceptual se establece exclusivamente a través del concepto de explicación. Toda explicación es, conceptualmente, unificación, y calificamos además una explicación como causal si además ella involucra lo que el concepto independiente de causa especifique. A esta versión de EU corresponderá en parte la explicación por subsunción modeloteórica que esbozaremos más adelante.

3) EU se reduce conceptualmente a EC. Esta alternativa excluye conceptualmente que pueda haber EC no unificadoras y supone por tanto la validez *a priori* de un principio que garantice que cuanto más acertadamente describimos la estructura causal del mundo más unificadas son nuestras descripciones. La existencia de un principio tal de uniformidad o simplicidad en la naturaleza *que valga a priori* es muy implausible (siempre y cuando no construyamos las relaciones causales como derivadas de las sistematizaciones unificadoras).

4) EU no se identifica con EC pero la presupone conceptualmente: toda EU es, conceptualmente, EC (aunque la inversa no es cierta). Esto es, las EU son, conceptualmente, EC *de cierto tipo es-*

pecífico: *x* U-explica *y* syss *x* C-explica *y* y además ... . De nuevo, obtenemos dos versiones según el concepto mismo de unificación dependa del de causa o no. Estaríamos en el primer caso, por ejemplo, si el concepto de unificación depende conceptualmente de la distinción proyectable/no proyectable (como en el caso de Kitcher) *y además* nuestro análisis de dicha distinción presupone, directa o indirectamente, el concepto de causa. En el segundo caso, se puede analizar el concepto de unificación sin apelar al de causa, pero no así el de explicación. Toda EU sería EC porque toda explicación es causal. Algún pasaje ocasional de Salmon parece estar en esta línea: «la tesis [de Kitcher] es que 'el debido a (*because*)' se deriva siempre del 'debido a (*because*)' de la explicación. Mi punto de vista es básicamente el contrario» (1994, p. 311). En ambos casos, es posible conceptualmente la existencia de explicaciones causales no unificatorias, por lo que no se presupone un principio de uniformidad de la naturaleza válido *a priori*. Lo que sí excluyen conceptualmente, sin embargo, es que pueda haber explicaciones (unificatorias) no causales. Diversos autores han ofrecido supuestos contraejemplos. Aunque muchos de ellos son controvertidos,[6] aquí nos basta observar que quienes defienden esta alternativa y además, como Salmon, proponen un análisis de la causalidad que admiten que no se aplica a teorías como la mecánica cuántica, y quizás tampoco a teorías psicológicas o sociológicas, deben aceptar que dichas teorías no son explicativas. En la medida en que pensemos, como es nuestro caso, que esta consecuencia es implausible, esta alternativa también lo será.

5) Ni toda EU es conceptualmente EC ni toda EC es conceptualmente EU, pero ambas comparten un núcleo conceptual común en virtud del cual hablamos en ambos casos de «explicación». En mi opinión, ésta es la alternativa de quienes rechacen dependencia conceptual en cualquiera de las dos direcciones pero acepten cierta complementariedad *de facto* en alguno de los sentidos extensionales que siguen, pues de otro modo, sin *ninguna* relación conceptual, la complementariedad extensional sería un hecho bruto sorprendente sin explicar. Es más, parecería que si no se acepta un núcleo conceptual común, ni siquiera habría motivo para indagar si ambos conceptos son compatibles o no lo son. Por otro lado, si se acepta,

6. Para una discusión de los mismos, cfr. Ruben (1990), cap. VII. Este autor presenta los suyos propios, principalmente los relacionados con los casos de explicaciones que involucran identidad de propiedades (p. ej., agua y H_2O, o temperatura y energía cinética media).

pero se rechaza alguna de las dependencias conceptuales fuertes anteriores, el problema entonces es determinar de modo preciso cuál es dicho núcleo (opciones como «esperabilidad» o «reducción a lo familiar» ya se han mostrado inadecuadas).

6) Solapamiento extensional fuerte. EC y EU dan *la misma* respuesta a *la misma* pregunta, entendiendo por «pregunta» *solamente* el tema junto con la clase de contraste, esto es, <*a*, *X*>. Así, dada una pregunta <*a*, *X*>, la misma respuesta es *a la vez* causal y unificadora. Esa pregunta forma parte de dos contextos <*a*, *X*, *R*> y <*a*, *X*, *R'*> que difieren sólo en la relación de relevancia explicativa, relevancia causal en un caso y unificativa en el otro. A pesar de ello, ocurre que la misma respuesta es relevante en los dos sentidos. Éste sería el caso de la explicación de la longitud de la sombra a partir de la longitud del mástil (y del ángulo de incidencia del sol), que puede ser vista como causal (Salmon) o como unificadora (Kitcher). Y también, la explicación de la órbita kepleriana de un planeta a partir de la fuerza de atracción del sol. Este solapamiento fuerte puede ser, a su vez: (6.i) *total*, si para toda pregunta, toda respuesta (explicativa) es a la vez EC y EU; (6.ii) *inclusivo*, si para toda pregunta, toda respuesta que es EC es también EU, o, en la otra dirección, toda respuesta que es EU es también EC; (6.iii) *parcial*, si hay preguntas con respuestas que son a la vez EC y EU. Nótese que estas relaciones extensionales *pueden* darse como consecuencia de las relaciones conceptuales anteriores, pues (1) y (3) implican (6.i), (2) y (4) implican (direcciones diferentes de) (6.ii) y (5) implica (6.iii). Pero alguien puede defender también que se dan estas relaciones extensionales sin las correspondientes intensionales, en cuyo caso serían hechos brutos sin explicar. Por otro lado, (6.iii) parece indiscutible,[7] y como veremos a continuación, (6.i) parece demasiado fuerte; la cuestión es si alguna dirección de (6.ii) es también admisible.

7) Solapamiento extensional débil. EC y EU dan respuestas *diferentes pero compatibles* a la misma pregunta, entendiendo «pregunta» como antes, esto es, <*a*, *X*>. En este caso, las relaciones de relevancia explicativa son tales que pueden no satisfacerse por la misma respuesta, cada respuesta es relevante respecto de una de las relaciones pero quizás no de acuerdo a la otra. A esta situación

7. «Estos dos tipos de comprensión se solapan con frecuencia» (Salmon, 1992, p. 39). El motivo que da, sin embargo, es discutible: «Cuando encontramos que el mismo mecanismo subyace a diversos tipos de fenómenos naturales, ello constituye *ipso facto* una unificación teórica» (*ibid.*); la cuestión es *por qué* cabe esperar que la explicación causal encuentre *los mismos* mecanismos causales operando tras diferentes tipos de fenómenos.

parece corresponder el ejemplo de Salmon del globo en el avión: ¿por qué al despegar un avión un globo relleno con helio se mueve hacia adelante en lugar de hacia atrás? Una explicación es que al acelerar el avión, la pared trasera «empuja» las moléculas hacia adelante creando un gradiente de presión que impulsa el globo. Otra explicación invoca el Principio de Equivalencia de Einstein, según el cual la aceleración es equivalente a un campo gravitacional, y puesto que los globos con helio tienden a ascender en el campo gravitacional terrestre, en el avión se comportan igual que si se creara un campo gravitacional en el punto donde comienza a acelerarse al avión, «ascendiendo» en el campo, esto es, moviéndose hacia adelante.[8] La primera explicación es mecánico causal, la segunda unificatoria, pero ambas son correctas. En términos de contextos explicativos, cada una responde adecuadamente a una misma pregunta de acuerdo con una diferente relación de relevancia explicativa. Ambas explicaciones son complementarias en el sentido de que ambas proporcionan información explicativa, pero información que es relevante desde diferentes «ángulos» (como a veces se dice, la causal sería «bottom-up» y la unificatoria «top-down», o «local» la primera y «global» la segunda). Esta relación de solapamiento débil puede ser, como antes la fuerte, (7.i) *total*, (7.ii) *inclusiva* o (7.iii) *parcial*.

8) Complementariedad sin solapamiento. EC y EU no se solapan, no dan nunca respuesta a las mismas preguntas, nunca explican los mismos fenómenos. Son por tanto excluyentes. Pero son complementarias en el sentido de que sus extensiones se complementan, todo fenómeno que tiene explicación tiene una explicación causal o unificatoria, y sólo una de ellas. En la medida en que aceptemos los casos anteriores de explicaciones causales y unificatorias de un mismo fenómeno (explicaciones iguales, como la de la órbita kepleriana de un planeta; o diferentes pero compatibles, como las del globo en el avión), esta última posibilidad es implausible. Sin embargo, sus defensores pueden replicar que tales casos están mal descritos. En el caso de la órbita del planeta, por ejemplo, en realidad no se explica el mismo fenómeno. La EC explica el suceso particular del movimiento del planeta. La EU no explica eso, sino el hecho general de que los planetas tengan tales órbitas, esto es, las leyes de Kepler. Así, un modo de defender esta alternativa sería postular que EC se aplica a hechos particulares y EU a hechos generales, a leyes. Pero esta defensa es poco plausible. Puesto que ambos explananda, hechos particulares y hechos generales, son dife-

8. Cfr. por ejemplo, Salmon (1989), p. 183.

rentes tipos de hechos, es seguro que la explicación no va a ser *exactamente igual*. Pero eso no quiere decir que el causalista no pueda dar un análisis de la explicación de leyes «en términos» causalistas y análogamente respecto del unificacionista para las explicaciones de hechos particulares.

Antes de hacer balance, conviene comentar brevemente la relación entre (6) y (7). Hemos visto que (6) y (7) describen tipos diferentes de solapamiento extensional, y también que cada uno puede ser de diferente grado. ¿Qué relaciones se dan entre las diversas posibilidades de (6) y (7)? Parece, en principio, que algunas excluirían otras. Así, por ejemplo, (7.iii) excluiría (6.i): si hay al menos un caso de respuestas *diferentes*, una que es EC y otra EU, a una misma pregunta, *entonces* no para toda pregunta sus respuestas son a la vez EC y EU. Sin embargo, para dar este paso hace falta aceptar *además* que la respuesta que es EC no es EU o que la que es EU no es EC. Pero eso es parte de lo que está en discusión. Podemos aceptar que en el caso del globo la explicación que apela al Principio de Equivalencia no es causal. Ello basta para descartar (6.i), pero no sólo como consecuencia de (7.iii). Y lo mismo sucede con la más débil (6.ii). Podemos aceptar que el caso del globo hace verdadero (7.iii), pero seguir aceptando alguna de las direcciones de (6.ii), por ejemplo, la que establece que toda explicación causal es también unificatoria. Si la explicación causal mediante el gradiente de presión es *también* unificatoria, entonces este ejemplo de (7.iii) no es un contraejemplo a (6.ii). Podemos aceptar pues (7.iii) y a la vez esa dirección de (6.ii), y por tanto (2). La cuestión que se plantea entonces es cómo puede haber dos EU *diferentes* de un mismo fenómeno. La respuesta es que la relación *R* de relevancia explicativa es un poco más compleja, no es simplemente «causalmente relevante para» o «unificatoriamente relevante para», sino que puede haber diferentes relaciones de relevancia todas ellas «unificatorias». En la medida en que sean realmente compatibles estarán relacionadas entre sí mediante algún tipo de «reducción interteórica». Pero entonces, aunque ambas explicaciones sean «correctas», una es mejor que otra. Eso es lo que pasaría en el caso de las explicaciones clásica y relativista del movimiento del globo, si la relativista reduce a la clásica.

Resumiremos ahora las principales consecuencias de esta revisión de las diferentes posibilidades de compatibilidad:

— Los conceptos EC y EU son diferentes, no se pueden reducir fuertemente entre sí identificando uno con otro de modo no trivial.

— Aunque posible, es poco iluminador aceptar relaciones extensionales brutas sin aceptar ningún tipo de relación conceptual.

— Es implausible que toda EU sea conceptualmente EC, al menos porque hay ejemplos no controvertidos de explicación cuya naturaleza causal es controvertida (salvo que trivialicemos la noción de causalidad). Casos como los de explicaciones mecánico-cuánticas o psicológicas, que Salmon descarta como causales; aunque parece que también considera controvertido que sean genuinas explicaciones. Casos como la explicación del movimiento del globo apelando al Principio de Equivalencia, que él parece aceptar como principio unificador no causal.

— No hay objeciones fuertes contra la idea de que toda EC es conceptualmente EU. Es poco plausible en la versión según la cual toda EC es EU porque el concepto de causa se analiza en términos de EU, lo cual supone la validez *a priori* de un principio de simplicidad en la naturaleza. Es más plausible en la versión según la cual toda EC es EU porque toda explicación es EU, y algunas son además EC por contener el ingrediente causal que el concepto independiente de causalidad especifica.

— La explicación de hechos particulares es de diferente tipo que la de hechos generales, pero ello no impide que se puedan analizar ambas «en términos de lo mismo» (causalidad o unificación).

— Puede haber explicaciones diferentes del mismo hecho e incluso del mismo tipo (causal o unificador), pues las relaciones de relevancia explicativa que determina el contexto pueden especificar más que el mero tipo causal o explicativo. En algunos casos, sin embargo, las explicaciones pueden «reducirse» unas a otras y, aunque ambas correctas, las reductoras son «mejores» que las reducidas.

Veamos ahora las líneas generales de una noción de *explicación* próxima a cierta versión de análisis modeloteórico de las teorías que parece especialmente adecuada de acuerdo con estas consecuencias. Esta noción de explicación va naturalmente asociada al programa modeloteórico de Suppes-Adams completado con algunos elementos de la concepción estructuralista, principalmente la distinción teórico/no teórico, la noción de teoría reticular y las condiciones de ligadura. Presentaremos brevemente este background metateórico,[9] antes de apuntar los líneas principales de esta noción de explicación.

9. Para una presentación un poco más detallada, cfr. Díez y Moulines (1997), cap. 10, n. 5; para una presentación completa, cfr. Balzer, Moulines y Sneed (1987).

3. **Presupuestos metateóricos**

(i) Análisis modeloteórico. Las teorías no se identifican mediante conjuntos de enunciados sino mediante conjuntos de modelos (aunque esos modelos son definidos mediante enunciados). Así, p. ej., para la identificación de la Mecánica Clásica serán relevantes los modelos del tipo <*P,t,s,m,f*> (*P* es un conjunto de partículas, p. ej., la tierra y la luna, *t* un intervalo temporal, *s* la función posición para las partículas en ese intervalo, *m* la función masa y *f* la función fuerza en ese intervalo) y otros definidos a partir de ellos.

(ii) *T*-teórico / *T*-no teórico. Los términos/conceptos de *T* son los términos (primitivos) que aparecen en la formulación de las leyes de *T*. Un término es *T*-no teórico si su denotación en el modelo se puede determinar (medir) sin usar las leyes de *T*, y es *T*-teórico en caso contrario, si todos sus procedimientos de determinación presuponen alguna ley de *T*. *masa* y *fuerza* son MC-teóricos, el resto MC-no teóricos.

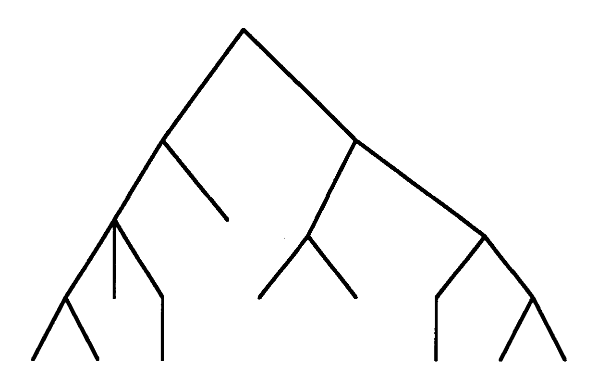
(iii) Fenómenos, base empírica. Los *modelos parciales* son los recortes *T*-no teóricos de los modelos de *T*; p. ej., en MC son del tipo <*P,t,s*>. Éstos son los modelos que representan los fenómenos particulares que explica la teoría, su base empírica. El conjunto de sistemas físicos concretos (p. ej., un niño columpiándose, la luna girando en torno a la tierra, ...) a los que se pretende aplicar la teoría es el *dominio de aplicaciones pretendidas I*.

(iv) Para dar cuenta de *I* la teoría impone diversos tipos de constricciones formales a los modelos. Las principales son las leyes. Las leyes que usan sólo términos de *T* definen un conjunto *M* de *modelos actuales*, p. ej., en MC las estructuras de tipo <*P,t,s,m,f*> que satisfacen además tales y cuales leyes. A veces algunas leyes usan conceptos de *T* y de otras teorías, son leyes «puente». Éstas también determinan un conjunto de modelos en *T*, el conjunto *L* de *vínculos*. El último tipo principal de constricciones se deben a requisitos de coherencia que *T* impone, p. ej., que cuando una partícula aparezca en dos modelos *m* le asigne la misma masa en ambos. El conjunto que expresa esta restricción sobre los modelos es el conjunto *C* de *ligaduras*. Así, podemos expresar el conjunto total de constricciones teóricas sobre los modelos mediante la terna *K*=<*M,C,L*>.

(v) Aserción empírica. Las constricciones formales de *K* tienen efectos a nivel *T*-no teórico. Si «juntamos» las de *M*, *C* y *L*, obtenemos el *contenido teórico CnT*, y si de los modelos de *CnT* recortamos su parte *T*- teórica entonces obtenemos el *contenido empírico* de *T*, *CnE*. La aserción empírica de la teoría es que los siste-

mas físicos de *I* cumplen las constricciones de *K*, esto es, que los sistemas de *I* están entre los de *CnE*.

(vi) Redes teóricas. No todas las constricciones formales están al mismo nivel, son igualmente centrales o básicas para *T*. Por ejemplo, en MC la explicación de cualquier fenómeno usa la Segunda Ley de Newton, pero no toda explicación usa la ley de Hooke (para un muelle sí, para un planeta no), o la de gravitación (para un planeta sí, para un proyectil también, para un muelle no), o la del rozamiento, etc. Las constricciones vienen estructuradas por estratos conformando una *red teórica*, empezando con las leyes (conjuntos de modelos) más generales y añadiendo después leyes más específicas en diversas direcciones para dar cuenta de fenómenos específicos (p. ej., en MC, una rama para fuerzas dependientes de la distancia, otra para fuerzas dependientes de la velocidad, etc.; una subrama de la primera para fuerzas dependientes directamente de la distancia, otra para las dependientes inversamente; etc.):



4. Explicación como subsunción

La noción central que recoge la idea de explicación dentro de este marco modeloteórico es la de *subsunción* (*embedding*). Veremos los principales rasgos de la misma,[10] el sentido en que se puede considerar más próxima al modelo unificacionista que al causalista, y la medida en que integra los principales hechos con que concluimos la discusión de la compatibilidad de ambos modelos. En principio, el aná-

10. Muchos de ellos son desarrollados por Bartelborth (1996, 1999, 2001).

lisis se aplica a los casos de explicación científica, esto es, en el contexto de teorías científicas desarrolladas, pero, aunque no lo haremos aquí, no sería difícil generalizarlo para otros casos precientíficos o *folk*.

La idea básica que se encuentra tras la noción de explicación como subsunción teórica es que explicar un fenómeno consiste en subsumir el fenómeno en un *patrón nómico*, esto es, en los términos modeloteóricos vistos, en subsumir el fenómeno en una rama de la red teórica de constricciones nómicas *K*. Explicamos, por ejemplo, el movimiento de la Luna en torno a la Tierra durante cierto intervalo *t*, subsumiendo el modelo parcial $y=\langle\{T,L\}, s, t\rangle$ en las constricciones teóricas *CnT* de cierta rama que comienza con la Segunda Ley de Newton, continúa con el principio de acción y reacción y prosigue hasta la Ley de Gravitación (más *links* y ligaduras). Aquí «subsumir» significa mostrar que en *CnT* hay modelos $\langle\{T,L\}, s, t, m, f\rangle$ que cumplen tales restricciones y que tienen a *y* como parte MC-no teórica. Así, explicamos ese fenómeno cinemático mostrando que, si la Tierra y la Luna tuviesen tales masas y estuvieran sometidas a tales fuerzas que se comportan con *m*, *s* y *t* del modo específico que tales constricciones determinan, entonces su movimiento debería ser el que de hecho es. Y análogamente con la explicación del movimiento de un objeto en el extremo de un muelle, o de un niño en un columpio, o de un cuerpo descendiendo en un plano inclinado con rozamiento, etc.

Es en este sentido, por tanto, en el que explicamos los fenómenos empíricos contenidos en el dominio *I* de aplicaciones propuestas subsumiéndolos en el núcleo nómico *K* de constricciones de la correspondiente rama de la red teórica. Nótese que la explicación de los fenómenos, *T*-no teóricamente identificados, la proporcionan los constituyentes *T*-teóricos de los modelos. Es *porque* en el sistema están «presentes» esos constituyentes *T*-teóricos comportándose con los *T*-no teóricos como las constricciones establecen, por lo que entendemos el comportamiento de la parte *T*-no teórica del sistema. Para que haya explicación por subsunción, por tanto, tiene que haber *extensión* o *ampliación teórica*. Aunque usualmente es lo que ocurre, no toda teoría genera extensión teórica, y aquellas que no la generen no podrán considerarse *explicativas*.[11]

11. Por ejemplo, la Cinemática galileana o la Astronomía kepleriana no lo son; los fenómenos *descritos* por estas teorías se explican a través de relaciones interteóricas reductivas que mantienen con otras (ver más adelante). Tampoco lo es la llamada *Measurement Theory* (cfr. Díez, 1998, 2000), pero en este caso la explicación de los fenómenos que describe no se obtiene por reducción sino mediante una relación interteórica más compleja (cfr. Díez, en prensa).

Este modelo de explicación de *fenómenos particulares* por subsunción teórica ampliativa en patrones nómicos se puede considerar próximo al modelo unificacionista en los siguientes respectos.

a) Los fenómenos no se explican «sueltos» sino «en grupos». Los elementos de I no son modelos parciales sueltos sino grupos de ellos, por ejemplo, la Luna dando vueltas a la Tierra, junto con Venus dando vueltas al Sol, junto con ..., o un niño en un columpio, junto con el péndulo de un reloj, junto con..., etc.

b) La explicación es tanto mejor cuanto mejor es el balance entre *aplicación* e *información*. La aplicación es tanto mayor cuanto mayor es la parte de *I* exitosa. La información es tanto mayor cuanto más restringe *K*, esto es, cuanto menor es *CnT*. Y estos dos parámetros compiten entre sí: se puede ampliar mucho la aplicación disminuyendo mucho la información (cuantos menos constricciones imponga *K* más fácil es que las aplicaciones de *I* estén efectivamente en *CnE*), y viceversa, hasta los casos extremos de aplicación total por información nula (tautológica) o información máxima con aplicación nula.[12]

c) La explicación es *integradora* u *orgánica*: (i) La explicación de unos fenómenos en una rama de la red está integrada con la de otros en otras ramas a través de la dependencia de ambas de las partes comunes de sus respectivas ramas, al menos las constricciones generales superiores (p. ej., la Segunda Ley de Newton) comunes a toda la red. (ii) La explicación conecta modelos entre sí a través de las condiciones de ligadura *C*. (iii) La explicación conecta con otras teorías a través de los vínculos interteóricos *L* que forman parte de las restricciones de *K*.

d) Dentro de una misma teoría, se produce progreso explicativo cuando fenómenos que requerían explicaciones diferentes se explican ahora del mismo modo. En términos de redes teóricas, cuando la red se simplifica *sin perder aplicatividad*. Para que ello represente un progreso efectivo se deben excluir casos en los que simplemente se «conjuntan» ramas diferentes. Para evitar esa situación, la red ha de ser tal que en ningún extremo terminal <*K*, *I*> de una rama sea posible dividir *K*, *I* respectivamente en *K*1, *K*2, *I*1, *I*2 tales que *I* es subsumible bajo *K* syss *I*1 es subsumible bajo *K*1 y *I*2 es subsumible bajo *K*2.[13]

12. Bartelborth (2001), pp. 9-11.
13. Esta condición modifica ligeramente otra de Bartelborth (2001), p. 10.

¿Es compatible esta versión de la explicación unificadora como subsunción ampliativa con la explicación causal? Sí que lo es, y lo es al modo de la opción (2) del segundo apartado. Esto es, la subsunción (unificadora) es conceptualmente más básica a la explicación que la causalidad. No nos preguntamos si hay operando mecanismos causales para poder decir que explicamos el fenómeno. Quizás los haya a veces, incluso pudiera haberlos siempre, pero en este marco eso no es lo conceptualmente relevante para hablar de explicación. Se explica un fenómeno cuando se subsume ampliativamente en la correspondiente rama de la red. La explicación será *además* causal si alguno de los constituyentes *T*-teóricos con los que ampliamos los modelos *T*-no teóricos tiene *potencia causal*, o alternativamente, si las constricciones-leyes de *K* son leyes causales. Muchas veces así es, como en la explicación mecánico-clásica de los fenómenos cinemáticos. Pero en algunos casos puede no ser así, o es cuando menos controvertido y ello no impide hablar de explicación. Algunos, por ejemplo, defienden que sólo las leyes de sucesión son causales, y las de coexistencia no. No podemos detenernos ahora en esta polémica, pero desde la perspectiva que hemos presentado la subsunción bajo cualquiera de ellas es explicativa independientemente de su eventual carácter causal. Quizás la explicación del movimiento del globo en el avión apelando al Principio de Equivalencia no sea causal, o la de la conducta en términos de deseos y creencias, o las mecánico-cuánticas, pero ello no obsta para que las califiquemos de explicaciones genuinas sin esperar a resolver la controversia, y lo hacemos porque independientemente de su estatuto causal también en ellas encontramos subsunción teórica.

¿Qué decir desde esta perspectiva subsuntivista de las explicaciones de *hechos generales*? En este fenómeno siempre está involucrado algún tipo de reducción interteórica, y la reducción interteórica se elucida, indirectamente, en términos de subsunción. La reducciones a la Mecánica Clásica de la cinemática galileana, las leyes de Kepler, la Mecánica del Sólido Rígido o la Termodinámica representan casos paradigmáticos de explicación por unificación. Ahora bien, el tipo de reducción no es siempre el mismo y hay que distinguir por lo general dos tipos de casos, según los hechos generales explicados sean a su vez explicativos o no lo sean. Las leyes galileanas del movimiento, o las leyes de Kepler son de este segundo tipo. Las leyes de la Mecánica del Sólido Rígido o las de la Termodinámica son del primero. La reducción de la cinemática galileana y la astronomía kepleriana a la MC es una re-

ducción exclusivamente a nivel empírico o aplicativo. Esta reducción consiste simplemente en que la teoría reductora incluye entre su dominio de aplicaciones pretendidas/modelos parciales que expresan las regularidades no explicativas de las teorías reducidas, modelos que son explicados mediante subsunción ampliativa en la teoría reductora. En este caso, pues, se trata de reducción explicativa pero sin *sustitución explicativa*. Un fenómeno *general* no explicado se explica en el marco de la reductora. Por otro lado, la reducción de la Mecánica del Sólido Rígido o la Termodinámica a la MC es una reducción teórica o de *sustitución explicativa*, pues estas teorías realizaban sus propias explicaciones por subsunción ampliativa de determinados fenómenos. La reducción explicativa en este caso consiste en lo siguiente: es posible correlacionar los modelos parciales y los completos de ambas teorías de modo tal que, el que ciertas aplicaciones de la reductora sean subsumibles en ciertas constricciones *K* de ella implica que las correspondientes aplicaciones de la reducida sean subsumibles en las correspondientes constricciones *K'* de ella. Esto es, una subsunción-explicación sustituye a otra. Y no sólo la sustituye, sino que, en cierto sentido, la explica. El que ciertas subsunciones-explicaciones (de la reductora) funcionen, explica que ciertas otras subsunciones-explicaciones (de la reducida) funcionen. Esta explicación entre explicaciones no es literalmente del mismo tipo que las subsunciones, pero se analiza en términos de subsunciones.

Por último, este modelo subsuntivista permite dar cuenta de manera natural de una característica ubicua de la explicación científica, al menos en las ciencias cuantitativas, a saber, su carácter *aproximativo*. Si las explicaciones cuantitativas pretendieran ser perfectamente precisas difícilmente podríamos aceptar ninguna. Es un aspecto central de ellas el que permitamos cierto margen de «error», y también de las cualitativas en la medida en que quepa hablar en ellas de «grados» de acierto. Este hecho se puede incorporar fácilmente en el modelo presentado. Al ser la relación de subsunción una relación modeloteórica, se puede «difuminar» o «emborronar» definiendo relaciones de aproximación entre modelos,[14] principalmente cuantitativos, pero se pueden aplicar igual a modelos cualitativos.

14. Cfr., por ejemplo, Balzer-Moulines-Sneed (1987), cap. VII.

5. Conclusión

La noción de explicación como subsunción teórica ampliativa expresa un sentido claro de «explicación» y acorde con los casos paradigmáticos de explicación científica. Este sentido es conceptualmente más básico que el de explicación causal y comparte muchos rasgos centrales del modelo de explicación como unificación. Este concepto, sin embargo, es compatible con el de explicación causal del modo más plausible de entre los que analizamos en la segunda sección. Toda explicación causal es explicación por subsunción. Y éste no es un hecho extensional bruto sino que se debe a vínculos conceptuales. Toda explicación causal es conceptualmente explicación por subsunción. El fundamento de este vínculo conceptual, sin embargo, no es una implausible dependencia conceptual de la noción de causa respecto de la de subsunción. Toda explicación es conceptualmente subsunción ampliativa, y será *además* explicación causal si involucra *además* los elementos que especifique el concepto independiente de causalidad.[15] Además, puede haber explicaciones diferentes de un mismo hecho. Cuando ello ocurre, eso se explica si se da alguna relación reductiva entre las dos explicaciones. Si no se da, tenemos dos explicaciones desconectadas en competencia. En cuanto a las explicaciones de hechos generales, ella no puede ser exactamente del mismo tipo que la de hechos particulares, pero se analiza naturalmente en términos de aquélla. El modelo, además, permite distinguir claramente dos tipos diferentes de explicación de hechos generales, y permite también incorporar de modo natural los aspectos aproximativos de la explicación.

Referencias bibliográficas

BALZER, W., MOULINES, C. U. y SNEED, J. D. (1987), *An Architectonic for Science. The Structuralist Program*, Reidel, Dordrecht.

BARTELBORTH, T. (1996), «Scientific Explanation», en BALZER, W. y MOULINES, C. U. (eds.), *Structuralist Theory of Science*, Walter de Gruyter, Berlín, pp. 23-43.

BARTELBORTH, T. (1999), «Coherence and Explanation», *Erkenntnis*, v. 50, pp. 209-224.

15. *Prima facie*, eso no quiere decir que, por motivos independientes de los tratados aquí, no se pueda acabar concluyendo tal dependencia, por ejemplo, si optamos por algún análisis epistémico de la causalidad.

BARTELBORTH, T. (2001), «Explanatory Unification», *Synthese* (en prensa).

DÍEZ, J. A. (1998), «Hacia una Teoría General de la Representación Científica», *Theoria*, v. 13, n. 31, pp. 113 -139.

DÍEZ, J. A. (2000), «Structuralist Analysis of Fundamental Measurement Theories», en BALZER, W., SNEED, J. D. y MOULINES, C. U. (eds.), *Structuralist Knowledge Representations. Paradigmatic Examples*, Poznan Studies in the Philosophy of Sciences and the Humanities, v. 75, Rodopi, Amsterdam, pp. 19-49.

DÍEZ, J. A., «A Program for the Individuation of Scientific Concepts» (en prensa), *Synthese*.

DÍEZ, J. A. y MOULINES, C. U. (1997), *Fundamentos de Filosofía de la Ciencia*, Ariel, Barcelona.

DOWE, P. (1992), «Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory», *Philosophy of Science*, v. 59, pp. 195-216.

DOWE, P. (1995), «Causality and Conserved Quantities: A Reply to Salmon», *Philosophy of Science*, v. 62, pp. 321-333.

FRAASSEN, B. VAN (1977), «The Pragmatics of Explanation», *American Philosophical Quarterly*, v. 14, pp. 1143-1150.

FRAASSEN, B. VAN (1980), *The Scientific Image*, Clarendon Press, Oxford.

FRIEDMAN, M. (1974), «Explanation and Scientific Understanding», *The Journal of Philosophy* v. 71, pp. 5-19.

KITCHER, PH. (1976), «Explanation, Conjunction and Unification», *The Journal of Philosophy*, v. 73, pp. 207-212.

KITCHER, PH. (1981), «Explanatory Unification», *Philosophy of Science*, v. 48, pp. 507-531.

KITCHER, PH. (1989), «Explanatory Unification and the Causal Structure of the World», en KITCHER, PH. y SALMON, W. C. (eds.), *Scientific Explanation*, University of Minnesota Press, Minneapolis, pp. 410-505.

KITCHER, PH. (1993), *The Advancement of Science*, Oxford University Press, Oxford.

KITCHER, PH. y SALMON, W. C. (1987), «Van Fraassen on Explanation», *The Journal of Philosophy*, v. 84, pp. 315-330.

KITCHER, PH. y SALMON, W. (eds.) (1989), *Scientific Explanation*, University of Minnesota Press, Minneapolis.

LEWIS, D. (1973), «Causation», *The Journal of Philosophy*, v. 70, pp. 556-567.

LEWIS, D. (1986), «Causal Explanation», en LEWIS, D., *Philosophical Papers II*, Oxford University Press, Oxford, pp. 214-240.

RUBEN, D. (1990), *Explaining Explanation*, Routledge, Nueva York.

SALMON, W. C. (1981), «Causality: Production and Propagation», en ASQUITH, P. D. y GIERE, R. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing (MI), pp. 49-69.

SALMON, W. C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton (NJ).

SALMON, W. C. (1989), «Four Decades of Scientific Explanation», en KITCHER, PH. y SALMON, W. (eds.), *Scientific Explanation*, University of Minnesota Press, Minneapolis, pp. 3-219.

SALMON, W. C. (1992), «Scientific Explanation», en SALMON, M. *et alt.*, *Introduction to the Philosophy Science*, Prentice Hall, Englewood Cliffs (NJ), pp. 7-41.

SALMON, W. C. (1994), «Causality Without Counterfactuals», *Philosophy of Science*, v. 61, pp. 297-312.

SALMON, W. C. (1997), «Causality and Explanation: A Reply to Two Critiques», *Philosophy of Science*, v. 64, pp. 461-477.

SALMON, W. C. (1998), *Causality and Explanation*, Oxford University Press, Nueva York.

SALMON, W. C. (2002*a*), «Explicación causal frente a no causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, pp. 97-115.

SALMON, W. C. (2002*b*), «Estructura de la explicación causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, pp. 141-159.

PARTE III

PROPUESTAS PARA LAS EXPLICACIONES CIENTÍFICAS

CAPÍTULO 4

EXPLICACIÓN CAUSAL FRENTE A NO CAUSAL[1]

WESLEY C. SALMON

Puede sorprenderles saber que, a comienzos del siglo XX, muchos filósofos y científicos creían que no había algo que fuera una explicación científica. Pensaban que la única función de la Ciencia era sistematizar nuestro conocimiento y predecir hechos futuros. Entre los más famosos defensores de esta concepción estaban los filósofos-científicos Pierre Duhem y Ernst Mach. De manera clara y sucinta lo señaló Karl Pearson (1911): «nadie cree ahora que la Ciencia *explique* cosa alguna; todos la vemos como una descripción breve, como una economía del pensar» (la cursiva es del autor). Quienes [entonces] creían que los fenómenos naturales *podían* ser explicados *no* buscaron sus explicaciones en el terreno de la Ciencia, sino más bien en los campos de la Teología y la Metafísica.

1. **Preludio a la «Concepción heredada»**

Esta situación cambió drásticamente hacia mitad de siglo. En los años treinta, Karl Popper defendía una clase de explicación deductiva similar a una idea que John Stuart Mill había articulado a mediados del siglo XIX. El trabajo de Popper, que apareció originalmente en alemán, fue publicado en inglés en 1959. En 1948, en sus *Estudios sobre la Lógica de la explicación*, Carl G. Hempel y Paul Oppenheim elaboraron un modelo concreto de explicación deductiva con gran precisión lógica. Más tarde, llegó a ser conocido como «modelo nomológico-deductivo». Tres aspectos importantes de ese

1. Ponencia presentada el día 8 de marzo de 2001 en las *Jornadas sobre Variedades de explicación científica*. La traducción ha sido realizada por Wenceslao J. González y José Fco. Martínez Solano.

trabajo requieren atención. (i) En ningún lugar de ese artículo se enuncia o se insinúa que toda explicación científica legítima se ajuste a ese modelo; por el contrario, los autores reconocen explícitamente la existencia de explicaciones inductivas o estadísticas en las Ciencias. Sin embargo, en 1948 no estaban preparados para ofrecer una teoría de las explicaciones de ese tipo. (ii) Casualmente, en el artículo de 1948, los autores identifican explicación deductiva con explicación causal. Hempel rechazó explícitamente esa identificación en 1965. (iii) El trabajo de 1948 tiene cuatro partes. En la parte I el modelo se presenta de manera informal. Se establecen entonces las siguientes condiciones preliminares de adecuación:

1. La explicación ha de ser un argumento deductivo válido.
2. El *explanans* debe contener esencialmente una ley general al menos.
3. El *explanans* ha de tener contenido empírico.
4. Las oraciones (*sentences*) que constituyen el *explanans* deben ser verdaderas.

El conjunto de premisas de un argumento explicativo se denomina «explanans»: aquello que explica (*which does the explaining*). La conclusión de ese argumento se llama «explanandum»: lo que es explicado.

La presentación formal y oficial de ese tipo de explicación se encuentra en la parte III. Muchos lectores han considerado que esos criterios expuestos en la parte I era todo lo que había. Esto es un error grave.

De manera extraña, en mi opinión, este artículo —que hace época— apenas atrajo la atención durante casi 10 años. Luego, a finales de los años cincuenta y primeros de los sesenta, hubo un elevado número de trabajos que se ocuparon del artículo de 1948 y de los temas que abordaba. Algunos de esos artículos suscitaron críticas altamente sofisticadas a la construcción formal de la parte III. El artículo de 1961 de Rolf Eberle, David Kaplan y Richard Montague es el más conocido. Trivializó el análisis de Hempel-Oppenheim mostrando, a grandes rasgos, que por lo mismo que se explica un hecho cualquiera se puede explicar cualquier otro hecho. En ese momento, cuando todavía estábamos encantados con la Lógica de predicados de primer orden, esas noticias producían excitación. Lo que pasó, de hecho, es que esos autores, para conseguir su resultado, explotaron las célebres paradojas de la implicación material. El problema podría ser resuelto mediante un cierto número de argucias

técnicas. En efecto, en el mismo número de la misma revista en la que apareció el trabajo de Eberle-Kaplan-Montague, David Kaplan publicó una solución satisfactoria del problema. Otros trabajos que aparecieron en esa misma época hicieron otro tipo de críticas (Sylvain Bromberger, 1962, es un ejemplo bien conocido), pero podían ser consideradas como «enmiendas amistosas» de autores completamente en sintonía con el modelo deductivo de Hempel-Oppenheim.

La principal crítica, empeñada en rechazar totalmente ese modelo, fue la expuesta por Michael Scriven (1958, 1959), que rechazó varios de los criterios básicos de adecuación sobre los que se basaba el modelo. Sostuvo, por ejemplo, que las explicaciones no necesitan incluir leyes de la Naturaleza —y, por tanto, rechazaba el aspecto nomológico del modelo nomológico-deductivo—; y mantuvo que las explicaciones de sucesos particulares comportan la identificación de sus causas. Aunque estoy de acuerdo con el requisito causal, no puedo aceptar la teoría de Scriven, porque nunca dio cuenta de la causalidad de una manera satisfactoria.

Antes de abordar el último trabajo de Hempel sobre la explicación [científica], debo hacer una observación extremadamente importante sobre todas las críticas, a saber, que *ninguna de ellas adoptó la posición de considerar a la explicación científica como intrínsecamente imposible*. Así, entre el comienzo y el ecuador del siglo XX, se produce un cambio de la actitud general, que pasó desde el rechazo casi completo a la casi completa aceptación. A principios de siglo, la posibilidad misma de explicación científica fue rechazada; medio siglo después aparece un consenso bastante amplio sobre la posibilidad, en principio, de la explicación científica. El problema estaba en encontrar caracterizaciones satisfactorias. En mi opinión, la razón principal de este giro fue la existencia de un modelo preciso y explícito de explicación. Esto les proporcionó a otros autores una idea concreta de lo que podía ser una explicación científica, y les dio un objetivo sobre el que disparar, si no estaban de acuerdo con el modelo que Hempel y Oppenheim habían articulado.

2. La «Concepción heredada»

Hempel no volvió sobre el problema de la explicación estadística hasta 1962, cuando publicó una exposición que él mismo encontró inadecuada. En su libro de 1965, *Aspectos de la explicación científica y otros ensayos sobre Filosofía de la Ciencia*, ofreció un marco

teórico general, que contenía cuatro tipos diferentes de explicación, que pronto se convirtió en una «Concepción heredada» (*received view*) entre los filósofos de la Ciencia. El marco [teórico] encuadraba el *planteamiento* de la explicación como «cobertura legal» (*covering-law*). (Hablar de «*modelo* de cobertura legal» es un error, porque se presentan cuatro modelos.) Toda explicación aceptable ha de satisfacer dos características decisivas. La primera es que debe ser o un argumento deductivo válido o un argumento inductivo correcto; esta ampliación de la primera condición de adecuación de Hempel-Oppenheim se requiere para dar cabida a las explicaciones estadístico-inductivas. La segunda es que debe contener una ley universal. Las leyes universales son generalizaciones universales y las leyes estadísticas son generalizaciones estadísticas. Las explicaciones se pueden también distinguir en términos de la clase de hechos que pretenden explicar, esto es, bien sean hechos particulares o bien regularidades universales. Los cuatro modelos son:

1. Explicaciones nomológico-deductivas de regularidades universales.

Ejemplo: La ley de conservación del momento se explica mediante la inferencia deductiva a partir de las tres leyes del movimiento de Newton.

2. Explicaciones nomológico-deductivas de hechos particulares.

Ejemplo: La vuelta del cometa Halley en 1759 fue explicada deduciéndola desde las leyes de gravitación y del movimiento de Newton, combinadas con las condiciones iniciales acerca de la configuración de los cuerpos en el Sistema Solar.

3. Explicaciones estadístico-deductivas de regularidades estadísticas.

Ejemplo: El hecho de que haya menos del 50 % de posibilidades de sacar el seis doble en 24 tiradas de un par de dados normales se deduce de las leyes de la probabilidad, la constitución simétrica de un dado legítimo y las leyes físicas relativas a la conducta de los sólidos simétricos.

4. Explicaciones estadístico-inductivas de hechos particulares.

Ejemplo: Aunque muchos colaboradores de Mary Smith contrajeron la gripe en el invierno de 2000-2001, ella no sufrió esa enfermedad porque recibió una inoculación, y en la mayoría (pero no en todos) los casos la vacuna confiere inmunidad ante la gripe.

Respecto de *todas* las explicaciones de hechos *particulares* (los tipos 2 y 4), Hempel dijo que «una explicación es un argumento que

tiene el efecto de explicar el hecho en cuestión, en cuanto que ha de esperarse a tenor de los hechos explicativos (*explanatory facts*)» (Hempel, 1962*a*). Esto significa que la conclusión de una explicación del tipo 2 se sigue deductivamente de sus premisas y que la conclusión de una explicación del tipo 4 tiene una alta probabilidad inductiva, en relación con las premisas del argumento inductivo. Las generalizaciones que constituyen las conclusiones de las explicaciones del tipo 1 y 3 han de ser aceptadas, en cuanto que se siguen deductivamente de leyes universales que ya están aceptadas. *La «Concepción heredada» de los años sesenta y setenta residía en que todas y cada una de las explicaciones legítimas se ajustaban a una u otra de las cuatro categorías señaladas antes.*

3. **Retos a la «Concepción heredada»**

Los retos a la «Concepción heredada» se presentaron a menudo por medio de contraejemplos ahora estándar (*now-standard*). Uno de los primeros fue presentado por Scriven (1959). La paresia es una forma de sífilis terciaria, que se presenta sólo en víctimas de la enfermedad que han atravesado las fases primaria, secundaria y latente sin tratamiento con penicilina. Si alguien llega a ser víctima de la paresia, la explicación se encuentra en esa secuencia de fases. Sin embargo, una persona que atraviesa las fases primaria, secundaria y latente de la sífilis sin tratamiento con penicilina sólo tiene un 25 % de posibilidades de desarrollar la paresia. La explicación es correcta, pero el resultado no sería el esperado, dada la información explicativa. Este ejemplo desafía el requisito de probabilidad alta para las explicaciones estadístico-inductivas.

Otro contraejemplo, citado con frecuencia, es el caso del barómetro y la tormenta. Si un barómetro registra una clara bajada de la presión [atmosférica], se puede inferir, con alta probabilidad, que sobrevendrá una tormenta. Sin embargo, el registro barométrico no explica la tormenta. Ambos se explican sobre la base de las condiciones atmosféricas, una causa común de los dos hechos. Esos dos ejemplos muestran que una probabilidad alta no es ni necesaria ni suficiente para explicaciones correctas. Este ejemplo apoya la exigencia de un ingrediente causal en la explicación.

Bromberger presentó otro famoso contraejemplo. Dado un mástil vertical de una determinada altura, que se sostiene en un terreno llano en un día soleado, la longitud de su sombra puede ser calculada sobre la base de la elevación del sol y de la ley de la pro-

pagación rectilínea de la luz. Esto constituye una explicación razonable de la longitud de la sombra. Sin embargo, para usar la misma clase de argumento, se puede calcular la altura del mástil desde la longitud de la sombra, pero esto no explica la altura del mástil. De nuevo se necesita un factor causal; el mástil causa la sombra, pero la longitud de la sombra no causa la altura del mástil.

Uno de mis contraejemplos favoritos está relacionado con el uso de contraceptivos orales; Henry Kyburg identificó originalmente el modelo. John Jones, un varón americano, evitó quedarse embarazado durante el año pasado porque consumió regularmente las píldoras anticonceptivas de su mujer, y todo hombre que tome regularmente contraceptivos orales evitará el embarazo. Nótese que es un argumento deductivo válido, y que, a efectos ilustrativos, cabe considerar como verdadero el supuesto de un hombre que tome píldoras anticonceptivas. Deseo resaltar que este ejemplo satisface todos los requisitos de Hempel para las explicaciones nomológico-deductivas. Obviamente, entraña un problema de relevancia, porque los seres humanos masculinos (a diferencia de los caballitos de mar) no pueden quedarse embarazados en ningún caso. La situación, en suma, es que los criterios de Hempel requieren la inclusión de *todos* los factores relevantes, pero no *sólo* de los hechos relevantes; es decir, *fracasan* para descartar factores irrelevantes.

Los factores irrelevantes invaden las explicaciones estadístico-inductivas. He aquí un contraejemplo complejo, que, a diferencia de los anteriores, posee una genuina significación científica. Supóngase que Robert Brown busca psicoterapia debido a un problema psicológico. A la vez que ha concluido el tratamiento, su problema se ha mitigado. Podría mostrar una fuerte inclinación a explicar su recuperación sobre la base del tratamiento, pero ésta no es una explicación necesariamente correcta. Es perfectamente posible que, para este problema particular, haya una alta proporción de casos de remisión espontánea, y que sea altamente probable que se pudiera haber recuperado incluso si no hubiera seguido la psicoterapia. En otras palabras, la probabilidad de recuperación tras la psicoterapia es alta, pero no es más alta que la proporción de casos por remisión espontánea. Si éste fuera el caso, la psicoterapia sería irrelevante desde el punto de vista explicativo; sin embargo, la «explicación» satisface los criterios de Hempel —incluida la condición de probabilidad alta— para la explicación estadístico-inductiva. Esta parte del contraejemplo muestra que una probabilidad alta no es *suficiente* para la explicación estadístico-inductiva.

En una variación sobre el mismo tema, podemos considerar a Betty Green, que asistió a sesiones de psicoterapia por un problema psicológico diferente. Supongamos que es baja la probabilidad de recuperación para pacientes que siguen el tratamiento; digamos que un 49 % (menos de la mitad). Sin embargo, ella se recupera. Supongamos, sin embargo, que la proporción de casos de recuperación espontánea en sus condiciones es muy baja; digamos que el 2 %. Dentro de esas circunstancias, sería razonable explicar su recuperación, al menos en parte, en términos del tratamiento. Este aspecto del contraejemplo muestra que no es *necesaria* una probabilidad alta para la explicación estadística. Aquí tenemos un tratamiento que es altamente relevante para el resultado, aun cuando ninguna probabilidad pueda ser considerada alta en términos del modelo propuesto por Hempel.

4. **La relevancia estadística**

Mi respuesta inicial, en los años sesenta, al problema de la relevancia en la explicación estadística fue resaltar el planteamiento según el cual, para la explicación estadística, el requisito central es la *relevancia estadística* en vez de la *alta probabilidad*. Esto llevó a la formulación de un modelo *de relevancia estadística* de la explicación estadística (Salmon, 1970). Nótese que el término «inductivo» no aparece en la denominación del modelo; esto significa que las explicaciones estadísticas *no son argumentos*. Según Hempel y la «Concepción heredada», todas las explicaciones legítimas son argumentos. Llamé a esto el *tercer dogma* del empirismo (siguiendo el famoso artículo de Quine «Dos dogmas del empirismo»), y lo rechazé frontalmente (Salmon, 1977).

De acuerdo con el modelo de relevancia estadística, se comienza con un caso concreto que pertenece a un grupo general de referencia. En uno de los célebres ejemplos de Hempel, tenemos gente que sufre una infección por estreptococo (cfr. Hempel, 1965, p. 381). En este caso, la persona —John Jones— se recobró rápidamente, y deseamos saber por qué. Mirando el ejemplo desde el punto de partida del modelo de relevancia estadística, hacemos una división relevante del grupo de referencia en dos subgrupos, a saber, aquellos que fueron tratados con penicilina y aquellos que no lo fueron. La probabilidad de una pronta recuperación es mucho mayor en el subgrupo tratado con penicilina que en el conjunto del grupo de referencia o en el subgrupo de aquellos que no

recibieron el tratamiento. En términos más simples, la división es relevante, porque el tratamiento con penicilina *marca la diferencia* en la probabilidad del resultado; esto es lo que significa *relevancia estadística*. No obstante, como señala Hempel, algunos enfermos tienen una infección que es resistente a la penicilina, y muchos de ellos no se recobrarán rápidamente, aun cuando sean tratados con penicilina. Esto constituye otra división relevante del grupo de referencia, en términos de resistencia o no a la penicilina. Un conocimiento adicional podría permitirnos hacer divisiones relevantes adicionales y, si ese conocimiento estuviera disponible, se necesitarían nuevas divisiones.

A tenor de las sugerencias del ejemplo precedente, podemos caracterizar una explicación de relevancia estadística como un conjunto de subgrupos compuesto *por todas y cada una* de las divisiones relevantes. La probabilidad del resultado en cuestión en el grupo general de referencia está dada; ésta es una *probabilidad anterior* (*prior probability*). Se asigna una probabilidad a cada subgrupo en cada división; éstas son las *probabilidades posteriores* (*posterior probabilities*). Se identifica entonces el miembro de la división a la que pertenece el individuo en cuestión; los factores explicativos son las propiedades que definen ese subgrupo. Por ejemplo, [el caso de] John Jones de Hempel pertenece al grupo de pacientes de estreptococo con infección no resistente a la penicilina, que fueron tratados con ella. Las relaciones de relevancia estadística llevan el peso de la explicación. En esta forma primitiva, las relaciones causales no entran en juego. Desde el principio reconocí que las relaciones causales son indispensables para la explicación; es un lugar común que las correlaciones estadísticas, por sí mismas, no constituyen relaciones causales (como resulta evidente en el ejemplo de la tormenta y el registro del barómetro). En esa época (en torno a 1970) pensé que las relaciones causales podrían ser analizadas por completo en términos de las relaciones de relevancia estadística, pero ahora considero que esa noción es insostenible.

Antes de dejar el tema de las relaciones de relevancia en general, sean estadísticas o de otro tipo, debo comentar la influyente teoría pragmática de la explicación de Bas van Fraassen (1980). Cualquiera que piense seriamente sobre la explicación, incluido [el propio] Hempel, reconoce que tiene muchas características pragmáticas. Esto no se cuestiona. Van Fraassen, sin embargo, creyó que la explicación podría ser entendida por completo sólo en términos de los aspectos pragmáticos. No puedo entrar en los detalles de la sofisticada teoría de van Fraassen, pero ésta también fracasa

en relación con la naturaleza de la relevancia explicativa. Philip Kitcher y yo hemos mostrado que el fallo de van Fraassen de poner restricciones objetivas sobre las relaciones de relevancia es lo que lleva a una posición en la que cualquier hecho puede explicar cualquier otro hecho (1987). Van Fraassen intentó superar las dificultades de las que estaba plagada la «concepción heredada», pero el esfuerzo, aunque valiente, no tuvo éxito.

5. **Explicaciones de generalizaciones**

Todos los contraejemplos que hemos encontrado hasta aquí tienen que ver con las explicaciones de hechos particulares. Son los casos en los que parece natural apelar a las causas. Sin embargo, Hempel y Oppenheim ofrecieron un contraejemplo a las explicaciones nomológico-deductivas de las regularidades universales (tipo 1), que se aplica igualmente a las explicaciones estadístico-deductivas de las regularidades estadísticas (tipo 3). Escribieron que «las leyes de Kepler, K, pueden ser combinadas con las leyes de Boyle, B, para [formar] una ley más fuerte, K-B.; pero la derivación de K a partir de la última no se considera una explicación de las regularidades enunciadas en las leyes de Kepler» (cfr. Hempel, C. G. y Oppenheim, P., 1948, nota 33). Confiesan su incapacidad para resolver esa dificultad. En un famoso artículo de 1974, Michael Friedman intentó superar el problema, pero su construcción resultó ser inadecuada. Se han hecho otros intentos, pero no tengo seguridad de que alguno de ellos haya funcionado. Si tuvieran éxito, el resultado podría ser un *modelo de unificación* (*unification model*) de explicación. No cuestiono la posibilidad de una teoría satisfactoria de esta clase; pienso que podría *complementar* en lugar de *entrar en conflicto* con la explicación causal.

En lo que queda de mis intervenciones, centraré la atención en las explicaciones de los hechos particulares. Dada la fuerza intuitiva de la idea según la cual apelar a las causas es lo apropiado, cuando deseamos explicar hechos particulares, necesitamos considerar las razones para resistirse a esta idea por parte de los defensores de la «Concepción heredada». Quizá se pueda encontrar una clara pista en el comentario que hace Hempel: «no está claro qué constructo preciso cabe ofrecer para la noción de factores 'que originan' (*bringing about*) un suceso determinado» (Hempel, C. G., 1965, p. 353). Sospecho que la dificultad de Hempel puede tener sus raíces en la famosa crítica de David Hume al concepto de «causa» en el si-

glo XVIII. El trabajo de Hume sobre la causalidad es el *locus classicus* de los debates modernos sobre el tema, y, dada la dificultad de los problemas que suscitó, debe ser nuestro punto de partida.

6. Hume sobre la causalidad

En su *Tratado de la Naturaleza humana*, Hume (1888, [1739-1740]) abordó [el tema de] la causalidad en detalle. Decepcionado por la recepción de su trabajo —puesto que dijo en su autobiografía: «*nació muerto desde la imprenta*» (1955 [1776], p. 4, cursiva de Hume)—, escribió un resumen anónimo (cfr. Hume, 1955 [1740]) en el que se centró en la causalidad como tema principal. Más tarde, en el *Ensayo sobre el entendimiento humano*, volvió al mismo tema. Debido a su naturaleza breve y enérgica, ese *abstract* sirve bien como punto de partida. Hume escribe:

> «Aquí está una bola de billar situada sobre la mesa, y otra bola se mueve hacia ella con rapidez. Chocan; la bola que, en principio, estaba parada ahora adquiere movimiento. Éste es un ejemplo tan perfecto de la relación de causa y efecto como ningún otro que se conozca por sensación o por reflexión. Por tanto, examinémoslo. Es evidente que las dos bolas se tocan entre sí antes de que se comunique el movimiento, y sin que exista pausa entre el choque y el movimiento. La *contigüidad* en el tiempo y el espacio es, por tanto, una circunstancia indispensable a la operación de todas las causas. Es evidente, asimismo, que el movimiento que estaba en la causa es anterior al movimiento que estaba en el efecto. La *prioridad* en el tiempo es, por tanto, otra circunstancia indispensable en toda causa. Pero esto no es todo. Ensayemos con cualesquiera otras bolas del mismo tipo y en una situación similar, y encontraremos siempre que el impulso de una produce el movimiento de la otra. Aquí, por tanto, hay una *tercera* circunstancia, a saber, la *unión constante* entre la causa y el efecto. Cada objeto como causa produce siempre algún objeto como efecto. Fuera de esas tres circunstancias de contigüidad, prioridad y conjunción constante, nada puedo descubrir en esa causa» (Hume, 1955 [1740], pp. 186-187).

El resultado de Hume es, desde luego, más famoso por sus omisiones que por su contenido. No consigue encontrar una relación *lógica* entre causas y efectos. Una vez que se tiene conocimiento de una causa, no podemos inferir, *sólo a partir de eso*, que se seguirá el efecto. Y al revés, teniendo conocimiento del efecto, no podemos inferir *sólo a partir de eso* cuál era la causa. En el célebre

ejemplo de las bolas de billar, no podía inferir lo que ocurriría cuando se encontraran las dos bolas. Quizá la segunda bola permanecerá parada, y la primera bola retrocederá en su movimiento y volverá a su punto de partida. Hume dice que podía imaginar esto sin [incurrir en] contradicción. ¿Se trata, simplemente, de una manifestación de la psicología de Hume? ¿Es genuinamente consistente lo que él imagina?

Podemos, de hecho, mostrar que es así. Dada la descripción de Hume de la situación anterior a la colisión, podemos añadir que la bola parada está asegurada firmemente a la mesa mediante un gran tornillo. Dadas estas circunstancias, la segunda bola se quedará parada después del choque. Si existe alguna duda sobre esto, podemos realizar rápidamente un experimento. Obviamente, todo lo que es real es lógicamente consistente. Hume no dijo nada sobre que la bola estuviera atornillada a la mesa, pero eso no importa. Consideren nuestra descripción de la situación y eliminen el enunciado (*statement*) sobre el tornillo. Es una simple cuestión de Lógica que un conjunto consistente de enunciados no puede volverse inconsistente meramente eliminando uno o más enunciados del conjunto. De modo que Hume tiene razón, como una cuestión de Lógica, y no meramente como una cuestión de Psicología. La misma consideración se aplica también en la dirección contraria. Podemos disponer una palanca oculta en la superficie de la mesa que le imprima movimiento a la segunda bola por un imperceptiblemente corto espacio de tiempo antes de llegar la primera bola. La misma palanca, cuidadosamente diseñada, parará el movimiento de la primera bola justo antes de que ocurra la colisión. Así, a partir del movimiento de la segunda bola, no podemos inferir la colisión con la primera bola. El argumento discurre *mutatus mutandus* como en el caso de inferir el efecto de la causa. Hasta aquí, el argumento de Hume es impecable.

Hume no sólo (correctamente) deja de encontrar una conexión lógica entre las causas y sus efectos; no logra asimismo encontrar una «conexión necesaria» entre las causas y los efectos, o un «poder secreto» por el que la causa produzca el efecto. Lo que de hecho reconoce es la «conjunción constante». Como una cuestión de «hábito» o «costumbre», la repetición de las observaciones de la causa seguida del efecto suscita en nosotros una *expectativa*, según la cual se seguirá el mismo efecto, dada la observación siguiente de la causa. La causalidad no es una característica del mundo físico conocido por la razón o por la observación empírica; es un producto de la *imaginación* humana. Según Hume, entonces, la causa-

lidad no es más que lo que hoy en día se denomina *respuesta condicionada* (*conditioned response*). Es como los perros de Pavlov. En su famoso experimento, Pavlov hacía sonar una campana justo antes de alimentar a sus animales de laboratorio. Después de haber repetido esta práctica un número de veces suficiente, los perros salivaban cuando sonaba la campana, expresando así su expectativa de [obtener] comida, aun cuando no se les presentaba comida alguna.

Me acuerdo de una famosa historia —no puedo garantizar su autenticidad— sobre Bertrand Russell y Alfred North Whitehead, los coautores de *Principia Mathematica*. Después de trasladarse Whitehead a la Universidad de Harvard, tras la I Guerra Mundial, Russell fue invitado a dar allí una conferencia. Eligió hablar sobre la Ética, y su discurso estuvo lleno de referencias a la Psicología científica de su época, incluyendo el condicionamiento psicológico y los perros de Pavlov. Al final de la conferencia, en el momento en que la audiencia iba a comenzar a aplaudir, Whitehead, que había presentado a Russell, se levantó de pronto y declaró: «verdaderamente, les digo a ustedes que el Bien es aquello que hace la boca agua». Podemos adaptar fácilmente el comentario de Whitehead: verdaderamente, Hume nos dice que *la causalidad es algo que nos hace la boca agua*.

7. La respuesta de Mackie a Hume

Si la concepción de Hume fuera correcta, implicaría que no puede haber algo como la causalidad en el mundo físico, al margen de las expectativas humanas. No podría haber habido causalidad antes de poblar el mundo las personas u otras criaturas inteligentes. La causalidad dejaría de existir si todas esas criaturas llegaran a extinguirse. Este punto de vista es difícil de aceptar. Estoy convencido, por ejemplo, de que el Gran Cañón del norte de Arizona fue producido —*causado*— por la erosión del río Colorado durante un periodo de muchos millones de años, mucho antes de haber existido en la Tierra seres humanos u otras formas de vida inteligentes.

No sorprende que muchos filósofos buscaran socavar la posición subjetiva de Hume. En un clásico del siglo XX, *El cemento del Universo*, J. L. Mackie (1974) pasa revista a alguno de los intentos más importantes. Creo que sus críticas a sus predecesores son en gran medida válidas. A continuación, elabora su propia teoría. Su objetivo era situar la causalidad «en los objetos» no «en la mente» (cfr. Mackie, 1974, pp. 1-2). Si su programa se hubiera llevado a cabo con éxito, debería haber dado como resultado una exposición rea-

lista —completamente objetiva— de la causalidad. Sin embargo, voy a argumentar que fracasa. Me centraré en tres puntos principales.

1. *Probabilidad lógica.* Mackie ofrece una detallada reconstrucción de la crítica humeana a la causalidad, pero sostiene que ha encontrado un defecto fundamental. Coincide en que Hume estaba en lo cierto al rechazar la noción según la cual, entre las causas y los efectos, pueden encontrarse conexiones lógicas *deductivas*; sin embargo, argumenta que Hume ignoró incomprensiblemente otra posiblidad, a saber, una relación de *probabilidad* lógica entre las causas y los efectos (*ibid.*, p. 15). Este camino no estaba abierto para Hume, porque la interpretación lógica de la probabilidad es esencialmente un producto de filósofos del siglo XX, como John Maynard Keynes (1921) y Rudolf Carnap (1950). No obstante, me parece que esa sugerencia es estéril, porque la introducción de las probabilidades lógicas incluye el uso de medidas arbitrarias de probabilidad, como pone de relieve el sistema de Carnap, altamente desarrollado (véase Salmon, 1967, 1969). Si alguien siguiera ese enfoque, tendría como resultado que la existencia o no de una relación causal en una situación particular estaría en el lenguaje elegido para describirla. Esa arbitrariedad lingüística es tan intolerable como la posición subjetiva de Hume.

2. *Regularidades causales.* Los escritos de Hume sobre causalidad no dejan claro si considera a la causa como condición suficiente de su efecto, como una condición necesaria, o como ambas. Para el presente tema, esto no importa. En cualquiera de esos supuestos, Hume está proponiendo la así llamada «posición de regularidad» (*regularity account*), esto es, o todos los *A* están seguidos de *B*, sólo los *A* están seguidos de *B*, o todos y cada uno de los *A* están seguidos de *B*. Cada uno es una generalización universal. Me producen una profunda insatisfacción esas teorías de la regularidad, porque es natural preguntar, en cualquiera de los casos, qué conexión hay entre los *A* y los *B* que sostenga la regularidad en cuestión.

Aunque Mackie no mantiene una teoría de la regularidad, ofrece un enfoque en el que juega un papel importante una regularidad bastante sofisticada, que comporta una compleja combinación de condiciones necesarias y suficientes. Distingue dos sentidos del término «causa». Según se usa de modo corriente, «causa» se refiere a una relación conocida como una condición INIS,[2] esto es, una

2. En el original es INUS: *Insufficient, Nonredundant, Unnecesessay, Sufficient.* (*N. del t.*)

parte **I**nsuficiente, **N**o redundante (esto es, necesaria), de una condición **I**nnecesaria y **S**uficiente (cfr. Mackie, 1974, p. 62).

De modo abstracto, la condición INIS es difícil de comprender, pero es fácilmente ilustrable. Supongan que se incendia un granero. Hay varias causas posibles, por ejemplo, un cigarrillo encendido lanzado por un fumador descuidado, la combustión espontánea del heno allí almacenado, un defecto en su instalación eléctrica, un acto deliberado de piromanía, ser alcanzado por un rayo, etc. Supongan, a título de ejemplo, que la causa real es el lanzamiento de un cigarrillo encendido por un fumador descuidado. Obviamente, ésta no es una condición necesaria, puesto que hemos enumerado varias causas posibles. Al mismo tiempo, el lanzamiento del cigarrillo encendido no es suficiente, por sí mismo, porque se deben dar otras circunstancias. El cigarrillo debe haber caído sobre algún material combustible tal como paja seca. La paja debe estar lo suficientemente cerca de otros materiales combustibles para que el fuego se propague. No debe haber habido alguien que apagara (*put out*) el fuego antes de su expansión. Y así, sucesivamente. Por consiguiente, el lanzamiento del cigarrillo encendido es una parte no redundante de un conjunto de condiciones que son suficientes para producir el fuego, pero ese conjunto no es necesario a la vista de otras posibles causas.

Mackie continúa con la caracterización de la «causa en el sentido filosófico» (*ibid.*, pp. 63-64), esto es, la *causa plena*, como una amplia disyunción de todas las causas suficientes, en la que cada una de las causas suficientes es una conjunción de todas las partes requeridas para conseguir la suficiencia. En el caso del granero, uno de los disyuntos es el lanzamiento del cigarrillo encendido, la presencia de paja seca, su proximidad a otros materiales combustibles, etc. Otro disyunto podría ser la combustión espontánea, esto es, el almacenaje de heno mojado, que se deja como está durante mucho tiempo; la acción bacteriana, el dejar de poner sal suficiente sobre el heno para combatir la acción bacteriana, etc. El resultado es una proposición en forma normal disyuntiva, que Mackie sostiene que es necesaria y suficiente para el efecto en cuestión. Esta condición necesaria y suficiente debe ser vista frente a la trayectoria de un campo causal de condiciones permanentes, tales como la presencia de oxígeno y la ausencia de asteroides que choquen con la Tierra en ese lugar determinado.

Concede Mackie la razonable posición según la cual elegir un factor, en lugar de otro, entre las condiciones INIS como la causa en el sentido corriente depende de todo tipo de consideraciones

pragmáticas y de contexto (*ibid.*, p. 73). Además, mediante la referencia de la causa plena a un campo causal, lleva a cabo la misma concesión respecto de la causa en el «sentido filosófico». Esto me parece a mí un reconocimiento de no haber tenido éxito para encontrar la causa «en los objetos». El campo es un rasgo del contexto, y depende de factores tales como nuestro conocimiento de fondo y nuestros intereses. Por otra parte, Mackie admite que el enunciado de las condiciones necesarias y suficientes es un «universal elíptico o con huecos (*gappy*)»: un enunciado que contiene espacios en blanco para la inserción de condiciones suficientes desconocidas hasta ahora (*ibid.*, p. 76). Esto no es un enunciado (*statement*); esto es la *forma lógica* de un enunciado que podría expresar una condición necesaria y suficiente del efecto en presencia del campo causal. Inmediatamente después de reconocer este punto, ofrece una detallada defensa de ese enunciado, pero su defensa descansa sobre su utilidad práctica, no sobre su objetividad independiente de la mente del sujeto (*ibid.*, pp. 67-75).

3. *Prioridad causal.* A Mackie le preocupa que su enfoque en términos de las condiciones INIS pueda ser víctima de la ausencia de asimetría causal, que él encontró en la teoría de Mill. Como en la teoría de Hume, en Mill la dirección de la causalidad está determinada únicamente por el orden temporal de los hechos. Coincido en que ésta es una característica indeseable en cualquier teoría de la causalidad. Para superar esta dificultad, Mackie presenta las nociones de *prioridad causal* y de *fijación causal* (*causal fixity*) (*ibid.*, pp. 178-183). La idea es que, en el *orden causal* de las cosas, las causas se pueden fijar (es decir, [quedar] completamente determinadas y permanentes) cuando los efectos no están fijados, pero los efectos no pueden ser fijados cuando sus causas no lo están. Un asunto distinto es si la prioridad temporal coincide siempre con la prioridad causal. Mackie propone que si *A* y *B* están directamente relacionados de modo causal, y si hay un periodo de tiempo en el que *A* está fijado pero no lo está *B*, entonces *A* debe ser causalmente anterior a *B*, y no a la inversa. Desafortunadamente, esta definición se desmorona por completo en algunas situaciones corrientes.

En relación con esto, tengo una pequeña historia que contar. Una tarde, hace muchos años, estábamos juntos Nancy Cartwright, John Earman y yo en un bar de Oberlin (Ohio), lugar de un buen *College* para estudiantes de Licenciatura que lleva el mismo nombre. El Departamento de Filosofía de Oberlin organizó un Congreso anual de Filosofía al que fuimos invitados los tres (Mackie asistió al congreso aquel año, pero él no estaba con nosotros en el bar esa no-

che). Mientras Earman estaba entretenido jugando al billar con un «tiburón» local, Cartwright y yo jugábamos a las máquinas. La máquina de bolas (*pinball machine*) es un aparato en el que el jugador lanza una pequeña bola de metal hacia la parte superior de una tabla inclinada. La bola desciende golpeando varios obstáculos, que se encienden y producen un tanto en ese particular juego. Durante el juego me di cuenta de repente de que teníamos ante nosotros un contraejemplo de la tesis de fijación de Mackie. Cuando la bola alcanza la parte superior y comienza a descender tenemos un hecho establecido *A*, y éste llevará inevitablemente a un resultado final *C*, esto es, la bola llega a pararse en un lugar determinado de la parte inferior de la tabla. El hecho *C* está fijado desde el momento en que la bola comienza su descenso desde la parte superior. No obstante, hay muchas trayectorias diferentes, *B*, por las que la bola puede desplazarse hasta la parte inferior; el camino concreto que la bola seguirá en un juego determinado no está establecido desde el principio. De acuerdo con la tesis de fijación de Mackie, entonces, la condición final *C* es previa causalmente a la trayectoria concreta *B* que la bola sigue para llegar a la parte inferior. Puesto que hay, muy a menudo, varios caminos diferentes para llegar al final, esta objeción es bastante general. Consecuentemente, la condición de fijación de Mackie no establece de modo adecuado la prioridad causal.

8. **Causalidad y explicación**

Como se dieron cuenta Hempel y Oppenheim en 1948, es esencial distinguir entre explicaciones de hechos particulares y explicaciones de regularidades generales. De modo intuitivo, parece razonable creer que podemos explicar las regularidades generales deduciéndolas a partir de leyes generales de nivel más alto. Las tres leyes del movimiento de Newton explican por qué se ha de conservar el momento lineal. Pongamos a un lado la dificultad de Hempel y Oppenheim citada en relación con las explicaciones de esta clase; no parece que sea insuperable.

Incluso intuiciones más fuertes conectan las explicaciones de hechos particulares con relaciones causales. Explicar por qué se estrella un avión *es* encontrar la causa o las causas del accidente. A pesar de esas últimas intuiciones, los principales enfoques filosóficos de la explicación científica han eludido de modo consciente la introducción de la causalidad en sus teorías. Creo que su prevención se debía a la ausencia de una teoría satisfactoria de la causalidad.

Entre los autores del siglo XX que tratan la causalidad, Mackie desempeña un papel preeminente. Obviamente, Mackie no ha tratado todas las teorías de la causalidad propuestas desde los tiempos de Hume, pero ha revisado algunas de las mejores candidatas (incluidas las correspondientes a los contrafácticos: Mackie, *ibid.*, cap. 2). Más aún, como parte esencial de su propia teoría, ha construido una concepción de la regularidad tan sofisticada que parece posible. Sin embargo, no consigue ofrecer una explicación de la causalidad completamente objetiva, libre de contexto. Sin negar los obvios aspectos pragmáticos de la explicación, sostengo que tanto la explicación como la causalidad necesitan una base enteramente objetiva. Así, creo que ha llegado la hora de proponer enfoques radicalmente diferentes de ambos conceptos, en especial de la causalidad. En mi segundo trabajo elaboraré lo que me parece el enfoque disponible más prometedor para la causalidad.[3] Mi meta será proporcionar una base sólida para una teoría causal de la explicación científica.

Referencias bibliográficas

BROMBERGER, S. (1962), «An Approach to Explanation», en BUTLER, R. S. (ed.), *Analytical Philosophy-Second Series*, Basil Blackwell, Oxford, pp. 72-105.

CARNAP, R. (1950), *Logical Foundations of Probability*, University of Chicago Press, Chicago.

EBERLE, R., KAPLAN, D. y MONTAGUE, R. (1961), «Hempel and Oppenheim on Explanation», *Philosophy of Science*, v. 28, pp. 418-428.

FRAASSEN, B. C. VAN (1980), *The Scientific Image*, Clarendon Press, Oxford.

FRIEDMAN, M. (1974), «Explanation and Scientific Understanding», *Journal of Philosophy*, v. 71, pp. 5-19.

HEMPEL, C. G. y OPPENHEIM, P. (1948), «Studies in the Logic of Explanation», *Philosophy of Science*, v. 15, pp. 135-175. [Reimpreso en HEMPEL, C. G. (1965), pp. 245-295].

HEMPEL, C. G. (1962), «Explanation in Science and in History», en COLODNY, R. G. (ed.), *Frontiers of Science and Philosophy,* University of Pittsburgh Press, Pittsburgh, pp. 7-34.

3. SALMON, W. C., «Estructura de la explicación causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, 2002, pp. 141-159. (*N. del t.*)

HEMPEL, C. G. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, Nueva York.

HUME, D. (1888 [1739-40]), *A Treatise of Human Nature*, edición de L. A. Selby-Bigge, Clarendon Press, Oxford.

HUME, D. (1955 [1740]), «Abstract of *A Treatise of Human Nature*», edición de C. W. Hendel (1955).

HUME, D. (1955 [1748]), *An Inquiry Concerning Human Understanding*, edición de C. W. Hendel, Bobbs-Merrill, Indianápolis. [La grafía moderna de «Enquiry» ha sido introducida por el editor.]

HUME, D. (1955 [1776]), «My Own Life», edición de C. W. Hendel, Bobbs-Merrill, Indianápolis.

KAPLAN, D. (1961), «Explanation Revisited», *Philosophy of Science*, v. 28, pp. 429-436.

KEYNES, J. M. (1921), *A Treatise on Probability*, Macmillan, Londres.

KITCHER, PH. y SALMON, W. C. (1987), «Van Fraassen on Explanation», *Journal of Philosophy*, v. 84, pp. 315-330.

KYBURG, H. E. Jr. (1965), «Comment», *Philosophy of Science*, v. 32, pp. 147-151.

MACKIE, J. L. (1974), *The Cement of the Universe*, Clarendon Press, Oxford.

MILL, J. S. (1874), *A System of Logic. Ratiocinative and Inductive*, (8.ª ed.), Harper, Nueva York.

PEARSON, K. (1911), *The Grammar of Science* (3.ª ed.), Meridian, Nueva York.

POPPER, K. R. (1935), *Logik der Forschung*, Springer, Viena.

POPPER, K. R. (1959), *The Logic of Scientific Discovery*, Basic Books, Nueva York. [Vers. ingl. de Popper, 1935.]

SALMON, W. C. (1967), «Carnap's Inductive Logic», *Journal of Philosophy*, v. 64, pp. 725-739.

SALMON, W. C. (1969), «Partial Entailment as a Basis for Inductive Logic», en RESCHER, N. (ed.), *Essays in Honor of Carl G. Hempel*, Reidel, Dordrecht, pp. 47-82.

SALMON, W. C. (1970), *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh.

SALMON, W. C. (1977), «A Third Dogma of Empiricism», en BUTTS, R. y HINTIKKA, J. (eds.), *Basic Problems in Methodology and Linguistics*, Reidel, Dordrecht, pp. 149-166. [Reimpreso en Salmon, 1998.]

SALMON, W. C. (1998), *Causality and Explanation*, Oxford University Press, N. York.

SCRIVEN, M. (1958), «Definitions, Explanations, and Theories», en FEIGL, H. ET AL. (eds.), *Minnesota Studies in the Philosophy of*

Science II, University of Minnesota Press, Mineapolis, pp. 99-195.

SCRIVEN, M. (1959), «Explanation and Prediction in Evolutionary Theory», *Science*, v. 30, pp. 477-482.

WHITEHEAD, A. N. y RUSSELL, B. (1910-1913), *Principia Mathematica*, 3 vols., Cambridge University Press, Cambridge.

CAPÍTULO 5

EXPLICACIÓN AXIOLÓGICA DE LAS ACCIONES CIENTÍFICAS

JAVIER ECHEVERRÍA

1. **Marco conceptual**

El debate sobre la explicación científica se ha centrado en el carácter causal o no causal del *explanans* y en la exigencia de que el *explanans* incluyera alguna ley científica a partir de la cual, incluyendo unas condiciones iniciales, se dedujera el *explanandum*. Se trataba de explicar el conocimiento científico, es decir:

A. Los hechos científicos, que habían de ser explicados por medio de leyes científicas.
B. Las leyes científicas (como la de Kepler), que habrían de ser explicadas por medio de leyes más generales (como las de Newton).
C. Las teorías científicas, que podrían ser explicadas subsumiéndolas en otras teorías más amplias y desarrolladas (reducción de teorías).

Así surgió el modelo deductivo-nomológico de explicación científica, propuesto por Hempel y Oppenheim (1948). Con anterioridad, Stuart Mill y Popper habían señalado el problema e indicado las principales vías para afrontarlo, pero fue Hempel quien propuso una formulación precisa y rigurosa del mismo. Su *covering law model* se convirtió en los años sesenta en uno de los grandes cánones de la *received view* en Filosofía de la Ciencia, sobre todo cuando lo amplió a la explicación estadística, tanto deductivo-nomológica como inductivo-nomológica (Hempel, 1965). El modelo hempeliano también se ha aplicado a la explicación teleológica. Braith-

waite y Nagel intentaron subsumir las explicaciones finalistas y teleológicas bajo el modelo de explicación deductiva, llegando a sugerir que el modelo deductivo-nomológico vale también para estos tipos de explicación científica (Nagel, 1977).

Tradicionalmente, la explicación teleológica ha tenido importancia en el ámbito de la Biología y se ha relacionado estrechamente con la explicación funcional.[1] El primer propósito de este artículo consiste en mostrar que existe un ámbito más general donde la explicación teleológica es frecuente, el de la actividad científica. En segundo lugar, trataremos de mostrar que varias teorías de la acción propuestas en la segunda mitad del siglo XX resultan insuficientes para explicar las acciones científicas. Por último, haremos nuestra propia propuesta de explicación axiológica de las acciones científicas, y en general de la actividad científica. Nuestro propósito consiste en mostrar que la Filosofía de la Ciencia, y por ende la teoría de la explicación, no debe limitarse al análisis y reconstrucción racional del conocimiento científico, sino que también ha de abarcar la actividad científica. Desde este punto de vista, tan importante es la explicación del conocimiento científico como la de las acciones científicas. Sin embargo, el modelo nomológico-deductivo no es válido a la hora de explicar las acciones científicas. Por ello propondremos al final un modelo diferente, que ilustra lo que llamamos *explicación axiológica de la actividad científica*.

2. Las acciones científicas

Las teorías disponibles de la acción se han basado en las creencias, deseos, estados mentales, preferencias e intenciones de los agentes, discrepando a la hora de atribuirles una función causal con respecto a las acciones. Se ha dado por supuesto que los agentes son individuos, conforme al individualismo metodológico, tan habitual en Epistemología y en Teoría de la decisión racional. Las dificultades para explicar ulteriormente las acciones sociales y colectivas han resultado insuperables, porque los deseos, las preferencias y las intenciones de los individuos no se suman para producir un deseo, preferencia o intención colectiva. En el caso de las acciones científicas intervienen siempre numerosos individuos, sobre todo en la Ciencia contemporánea. Por ello las *teorías atomistas de la acción*, basadas en el individualismo metodológico, resultan

1. Ver Salmon, W. C. (1989), pp. 111-116.

insuficientes para desarrollar una teoría de la acción científica. La existencia de *valores compartidos* por los científicos puede aportar una nueva clave explicativa de la actividad científica, como intentaremos mostrar a continuación. De esos valores compartidos se pueden desprender reglas (metodológicas, por ejemplo), normas, instrucciones y procedimientos compartidos, pero no leyes, por lo general, y mucho menos leyes naturales. Además, la relación entre los valores y las reglas o normas no es de deducción lógica. Por tanto, el modelo nomológico-deductivo muestra claras insuficiencias conceptuales a la hora de explicar las acciones científicas.

Aclaremos, en primer lugar, qué entendemos por acción científica, proporcionado una primera lista de *acciones científicas elementales*: calcular, demostrar, hacer operaciones matemáticas, observar, medir, hacer experimentos, leer (o escribir) una tabla de datos, etc. Por supuesto, hay acciones científicas más complejas, por ejemplo publicar un artículo, presentar un proyecto de investigación, refutar una hipótesis, elaborar una teoría, etc. Pero para lo que nos interesa podemos limitarnos a las acciones científicas más sencillas. Cabe decir que, para cada Ciencia, hay una serie de acciones científicas elementales que hay que saber hacer. La actividad científica tiene una fuerte componente técnica que incluye un *saber cómo*, no sólo un *saber que*. El conocimiento científico es teórico y práctico. A nuestro modo de ver, la Filosofía de la Ciencia ha mostrado muy poco interés por esa componente práctica de la Ciencia.

Resulta entonces que las acciones científicas siempre son *regladas*, es decir, han de ser llevadas a cabo respetando una serie de reglas que han sido enunciadas previamente. Si dichas reglas no son respetadas, la acción científica correspondiente (calcular, demostrar, hacer operaciones matemáticas, medir, observar, experimentar, etc.) no ha sido bien hecha, por lo que los resultados de dicha acción no son aceptados como científicos. Dicho de otra manera: para que el resultado de una acción científica sea tenido en cuenta, es un requisito imprescindible que la acción o acciones que dieron lugar a dicho resultado hayan sido bien hechas, es decir, conforme a una serie de reglas, normas y procedimientos. Así ocurre con los métodos prescritos y aceptados por la comunidad científica. Al calcular, medir, experimentar, etc., los científicos no proceden conforme a creencias individuales, sino que siguen *reglas intersubjetivas que rigen sus acciones*. Obviamente, las acciones científicas no son las únicas reguladas entre las acciones humanas, pero conviene tener muy presente que están estrictamente regladas y no dependen de deseos e intenciones individuales en tanto acciones científicas.

Hay un segundo requisito muy característico de las acciones científicas, a saber: las acciones científicas llevadas a cabo por alguien han de poder ser *repetidas por otras personas*, en principio por cualquier ser humano que esté en condiciones de llevarlas a cabo. De hecho, buena parte de la actividad científica consiste en repetir una y otra vez acciones llevadas a cabo por otras personas, y a poder ser por sus pasos, es decir, conforme fueron realizadas previamente por otros. La repetibilidad (potencialmente universal) de las acciones científicas es la característica más definitoria de la actividad científica, y no sólo en el caso de las Ciencias físico-naturales, sino también en Lógica, Matemáticas y Ciencias formales. Inferir y hacer silogismos son acciones, al igual que demostrar un teorema de Euclides o realizar un cálculo aritmético, algebraico o de otro tipo. Lo característico de la acción científica estriba en que dichas acciones siempre son repetibles (al menos en principio) y, además, son repetibles paso a paso, en todos y cada uno de sus detalles. De hecho, los científicos suelen comprobar la concordancia de sus datos intermedios, no sólo de los resultados finales. La Metodología científica se basa en esta posibilidad de repetir secuencias de acciones que dan los mismos resultados y siguen el mismo orden en que fueron hechas por otros. Al *actuar igual* los científicos producen los mismos fenómenos y los mismos datos, o aproximadamente los mismos. No estamos ante acciones colectivas, sino ante lo que podríamos denominar *acciones intersubjetivas*. Independientemente de sus deseos, estados mentales, preferencias e intenciones, las acciones científicas son iguales (o aproximadamente iguales) si producen los mismos resultados partiendo de condiciones similares y siguiendo un mismo orden de actuación.

La tercera propiedad general de las acciones científicas también tiene interés. Muchas acciones científicas requieren de instrumentos para ser llevadas a cabo. Las acciones científicas también se caracterizan por la posibilidad de que esos instrumentos sean operados por cualquier agente humano, siempre que éste posea unos conocimientos básicos sobre cómo usarlos, es decir, siempre que haya sido *formado como científico*. Los instrumentos automatizados evidencian de nuevo el carácter intersubjetivo de la mayoría de las acciones científicas elementales. Operados por diferentes personas, dichos instrumentos han de producir resultados similares y, en el caso de las Ciencias exactas, resultados exactamente iguales. Además, los resultados de las acciones elementales han de ser independientes de los instrumentos con los que se realizan. De no ser así, se definen con precisión los posibles sesgos e imprecisiones induci-

das por la utilización de unos u otros instrumentos. Este ideal no siempre es fácil de lograr, pero en el caso de las Ciencias formales resulta factible, al igual que en Física o Química. Por ello es importante tener en cuenta que muchas acciones científicas elementales, cada vez más, no son realizadas por agentes intencionales, sino por máquinas. Otro argumento más para considerar inadecuadas las teorías intencionales e individualistas de la acción a la hora de analizar las acciones científicas.

Podríamos mencionar otras particularidades, pero las tres anteriores pueden bastar para lo que aquí queremos proponer, a saber: las acciones científicas elementales dependen muy poco del agente que las practica ni de los instrumentos concretos que se utilicen. Otros agentes con otros instrumentos deben obtener los mismos resultados, siempre que respeten las reglas prescritas para la ejecución de dichas acciones. Podemos resumir lo anterior diciendo que las acciones científicas son (o deberían ser) intersubjetivas. O también: son tanto más científicas cuanto menos subjetivas. El ideal leibniciano del *calculemus* se aplica en gran medida a las acciones científicas. Diremos incluso que el conocimiento científico es intersubjetivo porque las acciones científicas son intersubjetivas, no al revés. La posibilidad de repetir las mismas acciones obteniendo los mismos resultados (o muy parecidos: siempre hay unos márgenes de error en las acciones, en los agentes y en los instrumentos) es la base praxiológica de la intersubjetividad científica.

Lo anterior no implica negar que haya diferencias en las acciones científicas: la destreza, la habilidad, la práctica, etc., generan diferencias apreciables entre un investigador avezado y un estudiante que aprende en un laboratorio de prácticas, o que se inicia en la resolución de cualquier problema matemático. Pese a ello, los resultados han de ser los mismos, aunque unos los obtengan con mayor facilidad, rapidez, precisión o acierto. El conocimiento científico puede ser compartido por los seres humanos porque las acciones científicas son intersubjetivas, o si se prefiere, porque no dependen del agente concreto ni de la máquina que las lleve a cabo. La *épisteme* científica se basa en la concordancia de acciones y resultados. Por eso no es *dóxa*. Cuando dos científicos anuncian cómputos, observaciones, mediciones o resultados experimentales distintos se produce un conflicto práctico, que tiene que ser resuelto haciendo concordar las dos acciones, y en particular sus resultados. La elaboración del conocimiento científico (hipótesis, teorías, leyes, etc.) es posterior a esa concordancia básica entre las acciones y sus resultados. Por ello es preciso explicar primero las acciones científicas.

Por tanto, la explicación científica no debe reducirse a la explicación del conocimiento científico, como ocurría en el modelo nomológico-deductivo de Hempel. Previamente hay que explicar las acciones que posibilitan dicho conocimiento.

Valga el ejemplo célebre de Hume a la hora de criticar la noción de causalidad.[2] Si una bola de billar choca con otra en una mesa podrá discutirse si el movimiento de la primera es la causa del movimiento de la segunda. Lo que nunca debe olvidarse, y menos en ese ejemplo, es que el movimiento de la primera es resultado de una acción previa, a saber, el golpe que un jugador le da con el taco de billar a la primera bola. La acción de golpear es la causa del primer movimiento, así como del choque ulterior (si la dirección en la que se golpeó fue atinada) e incluso del recorrido de ambas bolas, en función de la fuerza impresa en el golpe. Dos buenos jugadores de billar han de ser capaces de reproducir con gran exactitud sus respectivas acciones a la hora de intentar hacer una carambola. En lugar de problematizar la noción de causalidad a partir del momento del choque, Hume hubiera procedido mejor analizando la posible causalidad del acto de golpear con respecto al movimiento ulterior, tanto de la primera como de la segunda bola. Es un buen ejemplo de explicación teleológica. Lo que explica los movimientos de las bolas es el objetivo (logrado o no) de hacer carambola. Para lograr ese objetivo es preciso calcular la dirección, el impulso, el efecto, la colocación ulterior de las bolas, etc. De manera similar, entendemos que lo primero que hay que analizar, y en su caso explicar, son las acciones científicas, al igual que las acciones de los jugadores en la mesa de billar. El conocimiento científico no se logra sin acciones humanas previas: observaciones, mediciones, experimentos, cálculos, demostraciones, etc. Por ende, antes de explicar el conocimiento (hechos, leyes, teorías) hay que explicar las acciones. El debate sobre la explicación causal ha de empezar analizando las causas (motivos, razones, intereses, etc.) que provocan las acciones científicas.

Hasta ahora nos hemos referido a las acciones científicas elementales, no a las más complejas. Las peculiaridades individuales de los agentes pueden ser muy relevantes en este segundo tipo de acciones. Aun así, para que haya Ciencia ha de ser posible una práctica común que genere los mismos resultados o resultados muy se-

2. HUME, D., *Investigación sobre el conocimiento humano*, secc. VII, part. I, Alianza Editorial, Madrid, 1980, p. 87. Ejemplo comentado por Salmon en el primer trabajo del presente volumen, cfr. Salmon, W. C. (2002).

mejantes. La evaluación de las acciones científicas depende de su repetibilidad por otras personas. Valorar a fondo un experimento, una observación, una medición, un cálculo o una demostración implica en último término repetir paso a paso dichas acciones, para comprobar si hay algún error, es decir, para *comprobar que la acción está bien hecha*. Otro tanto ocurre al evaluar un texto: hay que leerlo con atención, paso a paso, repitiendo mentalmente el acto de escribirlo. Y los ejemplos podrían ser muchos más. Concluiremos por tanto que la Ciencia no existiría sin este requisito práctico: la *repetibilidad de las acciones científicas*.

Éste es uno de los argumentos que nos lleva a descartar muchas teorías de la acción, en la medida en que tratan de explicar las acciones y sus resultados en función de las creencias, deseos, preferencias e intenciones individuales de los agentes. Contrariamente a ellos, hay que adoptar una teoría de la acción donde la intencionalidad o las creencias subjetivas no sean más que un factor a la hora de evaluar, justificar y explicar las acciones, uno entre muchos. Las acciones científicas no se caracterizan por la libertad de elección de los agentes, sino por la gran cantidad de constricciones que sus acciones tienen. El ejemplo de las Matemáticas es suficientemente ilustrativo al respecto. Si dos matemáticos intentan resolver un sistema de ecuaciones, los resultados que obtengan no pueden depender de sus creencias e intenciones. Para que sus cálculos y deducciones sean considerados como resultados matemáticos, han de coincidir entre sí, independientemente de quién los haya obtenido. E incluso habrían de coincidir todos y cada uno de los pasos intermedios. Otro tanto ocurre en Física, y en general en las Ciencias duras. A continuación, la interpretación de los resultados puede ser distinta, entrándose en los debates hermenéuticos. Mas, por lo que respecta a la práctica científica, la repetibilidad de las acciones constituye una base común que, a nuestro modo de ver, es más sólida que la base empírica postulada por los positivistas.

Resumiendo: la intersubjetividad de las acciones científicas y la concordancia potencial de sus resultados es la base sobre la que se asienta la Ciencia. Una cosa es la correspondencia entre los enunciados científicos y el mundo real y otra muy distinta, y previa, la concordancia entre los resultados de las acciones científicas. Los científicos coinciden entre sí al actuar, aunque luego difieran al interpretar los resultados. Para la Filosofía de la Ciencia, el estudio y la explicación de las acciones científicas es previo al del conocimiento científico.

3. Explicar acciones y explicar hechos

Lo apuntado en el apartado anterior tiene consecuencias para la teoría de la explicación racional, porque entonces *las acciones científicas son previas a los hechos*. En algunos casos, los hechos científicos podrán ser explicados conforme a los modelos propuestos por Hempel y sus seguidores, e incluso podrá llegarse a explicaciones causales. En el caso de las acciones, es preciso apelar a otro tipo de explicación. En cualquier caso, los hechos científicos son resultados de acciones científicas, con la peculiaridad antes señalada de que dichas acciones son repetibles por diversos sujetos. Esa repetibilidad (reglada) de las acciones está a la base de la intersubjetividad de la Ciencia.

Si ello es así, difícilmente podemos seguir aceptando una teoría de la acción científica basada en el individualismo metodológico, es decir, en la relación causal entre las creencias, deseos, etc., de los agentes y sus acciones, y en particular entre las intenciones y los resultados de dichas acciones. Antes de teorizar y elaborar hipótesis y teorías a partir de datos, es preciso que los datos concuerden. Ello no sólo afecta al debatido problema de la adecuación empírica, que también, sino ante todo a la concordancia entre los sistemas de acciones que dan lugar a esos datos. Si aceptamos que, en el caso de la Ciencia, la acción es previa a los hechos y las teorías, el problema filosófico consiste en explicar la intersubjetividad, comunicabilidad y repetibilidad de las acciones científicas. La teoría de la explicación experimenta con ello un giro importante, al centrarse en las acciones intersubjetivas, antes que en hechos o leyes.

A partir de aquí, se entiende que el problema de la explicación teleológica resurja con fuerza, en la medida en que alguien pretenda hacer una teoría de la acción científica basada en los objetivos o metas de las acciones, es decir, en las finalidades. De hecho, buena parte de las teorías de la racionalidad científica en el siglo XX han estado basadas en explicaciones teleológicas. Para Popper, el objetivo de la Ciencia es la búsqueda de la verdad, para los positivistas el incremento de la adecuación empírica,[3] para Lakatos el descubrimiento de hechos nuevos y sorprendentes, para Laudan la resolución de problemas y para muchos sociólogos del conocimiento la realización de intereses y, en el mejor de los casos, el logro del consenso. Esas teorías de la racionalidad científica son insuficientes. No dispongo de espacio para justificar aquí por qué las explicaciones te-

3. Por ejemplo, Carnap y su grado de corroboración.

leológicas de la actividad científica no me parecen válidas,[4] pero, evocando a Rescher, es posible mencionar el argumento principal.

Se trata de contraponer la *racionalidad valorativa* a la *racionalidad instrumental*, entendiendo esta última como una modalidad de explicación teleológica de la racionalidad: el fin justifica y explica las acciones, y en su caso la elección de los medios. La racionalidad instrumental considera los fines como dados y explica las acciones en función de esos fines, teleológicamente. Sin embargo, desde el punto de vista de la racionalidad valorativa propuesta por Rescher y otros autores,[5] los fines de las acciones científicas son analizables y criticables en base al grado de satisfacción de determinados valores que el logro de los objetivos propuestos permitiría alcanzar. La racionalidad axiológica es previa a la racionalidad teleológica, y por ello la explicación axiológica es preferible a la explicación puramente teleológica de las acciones científicas. En lugar de reavivar la teoría de la explicación teleológica es preferible afrontar directamente el problema de la explicación axiológica de las acciones científicas. Es lo que haré en este texto, al proponer un modelo de análisis axiológico de las acciones científicas que mejore las explicaciones nomológico-deductivas y las teleológicas.

4. Análisis axiológico de las acciones científicas

En anteriores escritos he propuesto un primer modelo para el análisis y la explicación axiológica de las acciones científicas.[6] Dicho modelo no intenta proponer una teoría de la acción humana. Su ámbito de aplicación inicial se limita exclusivamente a las acciones científicas y tecnológicas, aunque en fases ulteriores podría corregirse para ampliarlo hacia una teoría general de la acción. Sus características más destacadas son las siguientes:

4.1. Se distinguen una serie de componentes A_z de una acción científica A. En Echeverría (2001*b*) distinguí nueve componentes, que ahora serán diez. En lugar de un *análisis monista* de las acciones científicas, por ejemplo en función de los objetivos, de las intenciones o de los resultados, como suele ser habitual en las teorías de la acción, propongo un *análisis pluralista*, basado en la distinción

4. Ver Echeverría, J. (1995) para una exposición más amplia al respecto.
5. Cfr. Rescher, N. (1988) y (1999), y también Agazzi, E. (1996).
6. Ver Echeverría, J. (2001*a*).

de diversas componentes, cada una de las cuales forman parte de la acción científica que se trata de analizar y explicar. Las intenciones y los resultados de las acciones científicas (mediciones, datos, etc.) son dos componentes a evaluar, no las únicas, como argumentaremos con detalle en este apartado.

4.2. Hecha esta precisión, la Teoría de la Acción está orientada axiológicamente, no teleológicamente, porque se centra en los diversos valores que permiten evaluar dichas componentes y explicar su incidencia en las acciones científicas. Cada una de esas componentes, por separado y conjuntamente, es evaluada en función de un subsistema de valores V_j. Dichos subsistemas, a su vez, son muy diversos: podemos considerar valores epistémicos, pero también técnicos, económicos, políticos, jurídicos, éticos, estéticos, etc. Por otra parte, cada subsistema incluye varios valores relevantes para unas u otras componentes de la acción A. Nuestro modelo de análisis de las acciones científicas también es pluralista en este sentido. En general, un subsistema V_i incluye una pluralidad de valores v_{ij}, algunos de los cuales pueden estar vinculados a valores de otros subsistemas. En conjunto, siempre hay un sistema de valores $V = \{V_i\}$ que permite evaluar una acción científica A, y ante todo evaluarla frente a otras acciones alternativas B, C, etc.

4.3. El modelo que proponemos incluye además factores de ponderación p_{ij} para cada uno de los valores v_{ij}. Dichos factores de ponderación pueden ser diversos según los subsistemas. No propugnamos un análisis atomista ni maximizador basado en una única componente de las acciones (por ejemplo los resultados) y en un valor primordial que habría que maximizar (por ejemplo la verosimilitud, o el grado de contrastación empírica). Por el contrario, proponemos un modelo *sistémico* basado en la noción de *satisfacción gradual*. Diremos que una componente concreta A_z de una acción científica satisface el valor v_{ij} en el grado g_z si y sólo si $v_{ij} (A_z) = g_z$, donde g_z es una magnitud que puede ser medida en escala cardinal, ordinal, de razón o de intervalos, dependiendo del tipo de valor v_{ij} del que nos estemos ocupando.

4.4. Por otra parte, la evaluación de las acciones científicas no puede ser concebida como un juicio puntual y definitivo, sino como un proceso que tiene lugar a lo largo del tiempo. Ello equivale a decir que los factores de ponderación pueden cambiar, e incluso pueden introducirse nuevos valores a lo largo del proceso de evalua-

ción, así como desaparecer otros que habían sido tenidos en cuenta inicialmente.

4.5. No se evalúan acciones aisladas, sino secuencias de acciones que se siguen de las anteriores. Denominamos *actividad científica* a ese conjunto de acciones que se desarrollan en el tiempo y se imbrican entre sí, influyéndose mutuamente. Como resultado final de dichas acciones pueden surgir hechos comprobados, leyes contrastadas y teorías propuestas. Pero esos constructos no llegan a ser considerados como científicos sin la evaluación favorable de todas y cada una de las acciones previas.

4.6. Partiendo de las premisas anteriores, es posible representar cada proceso de evaluación mediante una matriz de evaluación, $(p_{jk}.v_{jk}(A_i)\ (t))$, que puede ser analizada en cada instante t mediante diversos operadores matemáticos, el más sencillo de los cuales es el sumatorio (Â), y también el que se utiliza con mayor frecuencia.

4.7. Dicha matriz de evaluación es muy útil a la hora de sopesar y elegir entre acciones científicas alternativas, incluyendo conjuntos de acciones propuestas o realizadas, por ejemplo proyectos de investigación. La situación más habitual en los procesos de evaluación de la Ciencia y la Tecnología es así: en el caso de dos acciones alternativas, puede ser representada mediante la fórmula: $(p_{jk}.v_{jk}(A_i)\ (t)) \geq (p_{jk}.v_{jk}(B_i)\ (t))$, sin perjuicio de que pueda ocurrir que en el instante t_1 se prefiera la acción A a la acción B, o al menos algunas de sus componentes, y en un instante ulterior la situación haya cambiado porque las diversas componentes de la acción B hayan incrementado el nivel de satisfacción de los valores relevantes en el proceso de evaluación.

4.8. Esta representación formal de los procesos de evaluación no es normativa (de ella no se infiere lo que se debe hacer), pero tampoco se limita a ser descriptiva. La matriz que representa un proceso de evaluación concreto, una vez construida, permite intervenir en las acciones científicas incrementando el grado de satisfacción de unos valores o menguando el de algunos disvalores. Además, permite criticar los procesos de evaluación efectivos, bien mostrando que hay valores ocultos que no llegan a representarse explícitamente, bien indicando la insuficiencia de los subsistemas de valores intervinientes por comparación con otros procesos de evaluación que se hayan realizado previamente.

5. Componentes y valores de las acciones científicas

Distinguiremos inicialmente diez componentes de una acción científica A, y diremos que A = {A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8, A_9, A_{10}}:

El agente	A_1	Humano, artificial, institucional, colectivo...
Hace	A_2	Imagina, proyecta, propone, diseña, ejecuta, repite...
A	A_3	Objetos, cosas, personas, seres vivos, sistemas, entidades abstractas...
Con instrumentos	A_4	Herramientas, signos, lenguajes, máquinas...
Conforme a las reglas	A_5	Metodológicas, prescriptivas, operativas, normativas, proscriptivas...
En el escenario	A_6	Situación, contexto, lugar, época, cultura...
En condiciones	A_7	Físicas, cognitivas, económicas, sociopolíticas...
Con finalidades	A_8	Intenciones, metas, objetivos, problemas...
Con resultados	A_9	Posibles, plausibles, efectivos...
Y consecuencias	A_{10}	Previstas, imprevistas, colaterales, riesgos...

Examinemos con mayor detalle las diez componentes que hemos distinguido. En muchos procesos de evaluación científica se valoran los agentes que van a emprender tales o cuales acciones, exigiéndoles un nivel de competencia, prestigio, honestidad, conocimiento del asunto, especialización, etc. Cuando se evalúan agentes científicos (personas, equipos, instituciones, etc.) se aplican diversos criterios de valoración. Pero los criterios de evaluación de agentes no son los mismos que se usan para valorar los resultados

de sus propuestas y acciones, los instrumentos disponibles o las circunstancias que rodean las acciones propuestas (medios disponibles, financiación requerida, etc.). De ahí que, en general, no todos los valores del sistema V sean pertinentes para analizar la componente A_1, aunque algunos sí.

En segundo lugar, es importante tener en cuenta que las acciones científicas no sólo se evalúan cuando ya han sido realizadas, sino también antes de hacerlas. Lo más frecuente en Ciencia (y en Tecnología) es valorar proyectos de investigación, prototipos, propuestas, diseños de experimentos, etc. En muchas ocasiones la explicación de las acciones científicas realizadas ha de remitirse a las acciones propuestas que dan sentido a lo que luego se hizo.

En tercer lugar, no es lo mismo actuar científicamente sobre entidades vivas o inertes, o sobre animales o seres humanos, por ejemplo a la hora de hacer experimentos. Muchos problemas éticos y ecológicos surgen en función del tipo de objeto sobre el cual se va a actuar. Los criterios de valoración cambian en función de ello.

En cuarto lugar, también se evalúan los instrumentos precisos y los disponibles para llevar a cabo una determinada acción. En la Historia de la Ciencia es habitual tener que demorar algunas acciones en el tiempo por no disponer todavía de los instrumentos para llevarlas a cabo. Ello vale incluso para la acción de teorizar: Newton, por ejemplo, tuvo que elaborar su cálculo de fluxiones antes de desarrollar y llegar a exponer públicamente su teoría de la gravitación universal. Otro tanto cabe decir respecto a los instrumentos de observación (interferómetro de Michelson-Morley), de medida (microscopios electrónicos más precisos, telescopios más potentes, etc.), de cálculo (capacidad de cómputo o velocidad de los ordenadores), de experimentación (control de variables que afectan a los resultados del experimento) e incluso de demostración (véase el teorema de los cuatro colores). Obviamente, los criterios de evaluación de los instrumentos no sólo son epistémicos o técnicos, sino también económicos (inversión a realizar, coste del mantenimiento...), ecológicos (residuos, etc.) y de otros tipos. Los valores pertinentes para evaluar instrumentos científicos difieren netamente de los que se utilizan para valorar las tres componentes anteriores de la acción A. Algunos de ellos, como la precisión, pueden tener ponderaciones distintas a las que se utilizan, por ejemplo, para evaluar agentes científicos.

En quinto lugar, las acciones científicas siempre están regidas por reglas, y dichas reglas sintetizan y expresan diversos valores. Hay reglas metodológicas que llevan con seguridad a un resultado. Otras proscriben determinadas investigaciones, por haberse demos-

trado que es imposible resolver tales y cuales problemas. Algunas reglas son instrucciones de uso, otras consejos derivados de la experiencia. En general, el ámbito de las reglas que rigen la actividad científica es extraordinariamente complejo, por lo que merecería un estudio por separado, que aquí no pretendemos emprender. En cualquier caso, la actividad científica está regida por prescripciones y proscripciones, y algunas veces también por leyes, aunque éstas no son leyes naturales, sino jurídicas.

En sexto lugar, las acciones científicas siempre se desarrollan en una determinada situación o contexto físico, social, político o cultural. Como mínimo están situadas en el espacio y en el tiempo. Determinadas acciones son simplemente inviables en unas épocas o lugares, mientras que en otras pueden serlo. Otro tanto cabe decir del apoyo mayor o menor que unas acciones puedan recibir de las instituciones. Baste mencionar el *Analytical Engine* de Babbage para ilustrar lo que estamos diciendo.

En séptimo lugar, las acciones científicas siempre parten de unas condiciones iniciales, aunque sólo sea porque la mayoría de ellas requieren que otras acciones previas hayan sido realizadas con éxito. En función del estado de la cuestión o del estado del arte, como suele decirse, hay acciones que pueden ser consideradas como plausibles o descabelladas.

En octavo lugar, las acciones científicas, como cualesquiera otras acciones humanas, siempre tienen objetivos, metas e intenciones en base a las cuales se realizan. Pero las intenciones y los objetivos no bastan para explicar las acciones si, por ejemplo, no se cuenta con los recursos humanos adecuados (A_1) ni con los instrumentos precisos (A_3), o si no se siguen la reglas pertinentes, o si las consecuencias imprevistas son perniciosas, o si las condiciones iniciales son inadecuadas para el desarrollo de la acción intencional. La Ciencia no es cuestión de buenas intenciones. Por ello la axiología ha de tratar de valorar las intenciones y los objetivos, pero no explicar las acciones científicas en base a ellas.

En noveno lugar, pero sólo en noveno lugar, cabe analizar y evaluar finalmente los resultados de las acciones (datos, hipótesis, teorías, resultados experimentales, etc.). Las ocho componentes precedentes suelen ser evaluadas antes de que se emprendan efectivamente las acciones científicas. Por supuesto, los valores epistémicos tienen una enorme importancia al valorar la componente A_9, pero no sólo ellos. En la Ciencia contemporánea cada vez se valora más la utilidad social, política, económica, ecológica o militar de determinados resultados de la investigación científica. Valores como la

utilidad o la rentabilidad posible de los resultados de una investigación científica son altamente relevantes en los procesos de evaluación, tanto *a priori* como *a posteriori*. Por tanto, ni siquiera en este caso podemos restringirnos a los valores epistémicos a la hora de evaluar los resultados de las investigaciones.

En décimo lugar, no sólo importan los resultados directos, sino también sus consecuencias ulteriores, sean éstas efectivas o simplemente posibles. La fecundidad de una hipótesis o de una teoría es un criterio de valoración perfectamente racional, independientemente del grado de contrastación empírica de dicha hipótesis o teoría, como demostraron Lakatos y sus seguidores. Por tanto, hay que negar de nuevo la posibilidad de un análisis atomista de las acciones científicas. No sólo hay que tener en cuenta sus resultados, sino también sus consecuencias ulteriores. Las contraindicaciones de los productos farmacéuticos son un claro ejemplo de lo que estamos diciendo.

Entre las consecuencias, cabe mencionar explícitamente los riesgos de las acciones científicas, hasta el punto de que podrían ser considerados como una undécima componente de dichas acciones, aunque aquí los hayamos englobado en la décima para no multiplicar excesivamente las componentes a analizar. Muchas acciones científicas conllevan riesgos, y no sólo para los agentes que las hacen o los objetos que las sufren, sino también para los ámbitos donde se desarrollan. Valga un experimento nuclear como ejemplo, aunque podrían aportarse otros muchos. Uno de los graves problemas que ha de afrontar la Ciencia contemporánea es la evaluación de los riesgos que se derivan de las actividades científicas.[7]

Pudiera ocurrir que algunas de estas diez componentes tenga puntos de intersección con otras, e incluso pueda ser subsumida en alguna de ellas. Tampoco descartamos que, aparte de estas diez componentes, pudieran señalarse más, como ya hemos sugerido en el caso de A_{10}. En cualquier caso, este modelo de análisis nos muestra que las acciones científicas tienen una complejidad considerable, al menos si las contemplamos desde la perspectiva axiológica en la que nos hemos situado. A partir de las consideraciones anteriores, los criterios de evaluación efectivamente usados por los científicos (y por la sociedad) son muchos y diversos. Por ello afirmamos al principio la tesis del pluralismo axiológico frente al monismo, o tendencia a primar un valor y elaborar concepciones de la racionalidad científica en base a un único criterio.

7. Ver al respecto López Cerezo, J. A. y Luján, J. L. (2000).

Dicho pluralismo no implica afirmar una indiferencia entre los diversos valores. En la Ciencia en general, y en cada proceso de evaluación en particular, unos valores son más relevantes que otros. Por ello introducimos un factor de ponderación p_{ijk} para cada valor v_{ijk}. Además, el sistema V tiene una estructura, porque pueden distinguirse en V diversos subsistemas V_j, algunos de los cuales siempre son relevantes (como los valores epistémicos) y otros según los casos y las disciplinas científico-tecnológicas. En contribuciones previas hemos distinguido doce subsistemas de valores que, en mayor o menor grado, suelen ser tenidos en cuenta en los procesos de evaluación científica:[8] valores epistémicos, tecnológicos, económicos, ecológicos, políticos, jurídicos, sociales, militares, éticos, estéticos y religiosos, más lo que podríamos denominar *valores básicos*. No todos esos subsistemas de valores son relevantes para analizar una acción científica cualquiera, pero todos pueden serlo. En el caso de las acciones más complejas (proyecto Manhattan, exploración del espacio, proyecto genoma, etc.) casi todos los subsistemas intervienen. Por eso hay que partir de esos doce subsistemas para clasificar los valores que rigen o inciden en la actividad científica.[9]

Llegados a este punto, conviene explicitar algo que se desprende de las consideraciones precedentes y que resulta central en la metodología de análisis axiológico aquí propuesta. En general, los valores subyacentes a la actividad tecnocientífica no son analizables en función del significado de los términos que se utilicen para designarlos. En lugar de analizar el significado de los *términos axiológicos*, que puede variar según las componentes, proponemos una axiología basada en clases de modelos, siguiendo las líneas generales de la concepción semántica en Filosofía de la Ciencia.

Hechas estas precisiones, sigamos adelante. Puesto que los sistemas de valores V son del tipo $V = \{V_k\}$, siendo $V_k = \{v_{jk}\}$, el sistema de valores V es del tipo $V = \{v_{ijk}\}$, siendo k el subíndice de los subsistemas de valores, j del conjunto de valores de cada subsiste-

8. Cfr. Echeverría, J. (2001*b*).

9. Entre las diversas distinciones que se pueden proponer hay algunas de tipo formal, que son las más interesantes para el análisis axiológico que estamos proponiendo. Aparte de la distinción entre valores nucleares (o centrales) y valores orbitales (o periféricos), cabría distinguir clases de valores en los diversos subsistemas según las escalas de medida utilizadas (cardinales, ordinales, de intervalos y de razones). También cabría distinguir entre subsistemas intensionales y extensionales. Otras clasificaciones dependen de criterios de contenido y son las más tradicionales: por ejemplo la distinción entre valores internos y externos (Laudan) o entre valores subjetivos y objetivos (Putnam), así como la distinción clásica entre valores éticos, epistémicos, religiosos, económicos, etc., a la que también recurrimos.

ma e *i* el subíndice de las componentes de A para las que son relevantes los valores v_{ijk}. Pues bien, el análisis axiológico de las acciones científicas en el momento t_0 puede ser representado mediante la matriz de evaluación (v_{ijk} (A_i) (t_0)), que será comparada con las matrices de evaluación de otras acciones alternativas, (v_{ijk} (B_i) (t_0)), (v_{ijk} (C_i) (t_0)), etc. La elección de una acción como preferible a otra dependerá del algoritmo matemático que se elija para conjuntar valores que, en general, son heterogéneos, así como de las ponderaciones que se atribuyan (intersubjetivamente) a los diversos valores considerados. En el caso más sencillo, y por otra parte muy frecuente, se usa el sumatorio para agregar esas evaluaciones heterogéneas, de modo que la opción A es preferible a la B si y sólo si

$$\hat{A}\ (p_{ijk}.v_{ijk}\ (A_i)\ (t_0)) > \hat{A}\ (p_{ijk}.v_{ijk}\ (B_i)\ (t_0)).$$

Cuando no se evalúa únicamente una acción concreta, sino una secuencia de acciones es preciso considerar la variable tiempo. Otro tanto ocurre cuando una misma acción científica es valorada en dos instantes temporales distintos, lo cual es muy habitual, porque las acciones científicas se repiten una y otra vez, como ya indicamos al principio. En general, la matriz de evaluación M representa un proceso de evaluación de la acción A durante un intervalo de tiempo:

$$M(V)\ (t) = (p_{ijk}.v_{ijk}\ (A_i)\ (t))$$

Una acción científica A puede ser mejorada si cambia alguna de las componentes A_i, por ejemplo los instrumentos utilizados, de modo que la acción es más rápida, los resultados más precisos, los riesgos menores, etc. Todas estas *mejoras científicas* son fácilmente representables en la matriz de evaluación. En tales casos algunas de las casillas de la misma ven incrementada su magnitud, es decir, aumenta el grado de satisfacción del valor v_{jk} por parte de la componente A_i (o disminuye el grado de un disvalor). La práctica científica cotidiana suele porfiar por esas pequeñas mejoras, que son perfectamente representables mediante el instrumento formal M(V) (t). Conviene subrayar que los cambios de valores en la práctica científica también pueden ser fácilmente representados, y no sólo los cambios de ponderación de uno u otro valor. Cuando, tras haber valorado una innovación tecnocientífica en función de valores epistémicos y tecnológicos, pasamos a considerar otros aspectos, como su rentabilidad económica, sus riesgos ecológicos, su impacto social y jurídico o sus consecuencias éticas, estamos modificando el sistema

de valores V, y por ende la matriz de evaluación. En resumen, la introducción de los diversos subsistemas de valores modifica la matriz de evaluación, al igual que puede hacerlo la introducción de un nuevo criterio de valoración. En tanto instrumento puramente formal, las matrices de evaluación son adecuadas para representar los procesos de evaluación que efectivamente tienen lugar en la actividad científica. Obviamente, dentro de la matriz podemos distinguir submatrices para cada uno de los subsistemas efectivamente intervinientes en la evaluación. De la construcción de esas matrices se ocupa la *axiometría de la Ciencia*.

La *axiometría* que propugnamos está basada en la noción de *satisfacción* (total, parcial, gradual), no en la noción de maximización. En algunos casos será posible hablar de maximización en un sentido riguroso, pero en general no, habiendo de recurrir a indicadores del grado de satisfacción de un valor, que no siempre serán entidades métricas en un sentido estricto. La satisfacción de determinados valores hasta un cierto grado funciona como un requisito axiológico o condición *sine qua non*, no como una magnitud con la que se pueda operar aritméticamente introduciendo estrategias de maximización.

Dos observaciones más antes de terminar este apartado. En primer lugar: en la actividad tecnocientífica no se hacen valoraciones absolutas, sino relativas a otras acciones propuestas. Así, el resultado de una acción de teorización (una teoría, una demostración, una hipótesis, etc.) nunca es valorado *per se*, sino en relación a otras teorías, demostraciones o hipótesis propuestas. Esto es válido para las diez componentes de una acción tecnocientífica. El modelo de análisis que proponemos compara siempre dos acciones (o sus componentes), no versa sobre una acción aislada.

En segundo lugar, no optamos por la noción de preferencia, que tiene connotaciones subjetivistas, sino por un análisis meliorista (Dewey), donde la expresión «una acción tecnocientífica *A* (o una de sus componentes) es *mejor que* otra acción *A*» puede ser analizada en base a valores: ser mejor equivale a satisfacer en mayor grado valores pertinentes para evaluar dichas acciones (o componentes).

6. **Explicación axiológica de las acciones científicas**

Estamos en condiciones de afrontar la cuestión inicial de la explicación de las acciones científicas, no para resolverla, claro está, pero sí al menos para indicar la dirección en la que dicho problema

puede investigarse, partiendo de las consideraciones anteriores. Nuestra propuesta consiste en afirmar que *las acciones científicas y tecnológicas pueden ser explicadas en base al grado de satisfacción de un sistema de valores por parte de sus diversas componentes*. Dichos valores pueden ser subjetivos, pero en su mayor parte son intersubjetivos, es decir, valores compartidos por las comunidades científicas. En lugar de explicar esas acciones de manera atomista, conforme a las pautas del individualismo metodológico, optamos por un análisis axiológico, sistémico e intersubjetivo, basado en la distinción de diversas componentes de una acción, cada una de las cuales es evaluada en función de un subsistema de valores. Esta teoría de la explicación tiene la ventaja de ser gradual, al distinguir grados de explicabilidad de las acciones científicas. En la medida en que una acción satisface en alto grado diversos valores epistémicos, tecnológicos, económicos, etc., la realización efectiva de dicha acción está justificada en alto grado, y por ello es axiológicamente explicable que se realice esa acción, y no otra con grados inferiores de satisfacción de los valores v_{ijk}. Para que una observación, medición o experimento científico sea repetido, comprobado, precisado, etc., es requisito indispensable que dichas acciones satisfagan en alto grado los valores relevantes, es decir, por encima de un umbral mínimo de satisfacción. Si una acción A la llevan a cabo agentes prestigiosos, con instrumentos contrastados, en condiciones iniciales adecuadas, con objetivos precisos, etc., y si, además, dicha acción ofrece resultados relevantes (innovadores, precisos, útiles, rentables, etc.) y sus consecuencias y riesgos han sido evaluados en la medida en que ello es posible, diremos que esa acción es intersubjetivamente aceptable, o que es científica. Ello no implica afirmar que sus resultados sean verdaderos ni que el grado de eficiencia al hacerla no pueda incrementarse. Al sobrepasar los umbrales mínimos de satisfacción de los diversos valores relevantes v_{ijk}, la acción y sus resultados se convierten en un referente científico que otros repetirán e intentarán mejorar. Los avances científicos siempre se miden por el incremento del grado de satisfacción de algunos valores relevantes, $v_{ijk}(A_i) = g_{jk}$, siendo interpretable dicho incremento en función de un valor concreto j o de todo un subsistema de valores k. Este último caso es muy frecuente cuando una acción científica tiene en cuenta (o no) los valores ecológicos, o las normas jurídicas, o las consecuencias éticas, o simplemente los valores estéticos. Reducir el análisis y la explicación axiológica a los valores epistémicos y tecnológicos, sin tener en cuenta los restantes subsistemas, implica una simplificación que falsea el análisis y la explicación de las acciones científicas.

¿Es causal este tipo de explicación axiológica? En algunos casos puede serlo, pero en general no lo es, a no ser que entendamos los valores como causas formales o ideales regulativos. Por ejemplo, el incremento del grado de precisión (o de velocidad) de un instrumento de cómputo es causa de que dicho artefacto pueda resolver (o no) un problema científico. La disponibilidad de un determinado equipamiento científico es causa de que pueda afrontarse (o no) un determinado experimento o proyecto de investigación. Cuando las acciones científicas son propuestas, por ejemplo un proyecto de investigación, y son denegadas, hay que indicar las causas de la denegación. Pero el término «causa» no se entiende aquí en un sentido fisicalista, como en el caso del modelo hempeliano, sino en una acepción mucho más modesta, que tiene que ver con la noción leibniciana de requisito *sine qua non*, o de condición necesaria, pero no suficiente.

La tesis del pluralismo axiológico implica explicaciones de las acciones no monistas, sino plurales. Dentro de ese marco plural cabe hablar de *requisitos causales* (necesarios, pero no suficientes), más que de causas, en el sentido estándar del término. Pero este punto queda abierto al debate y a ulteriores publicaciones.

Referencias bibliográficas

AGAZZI, E. (1996), *El bien, el mal y la Ciencia. Las dimensiones éticas en la empresa científico-tecnológica*, Tecnos, Madrid.

BRAITHWAITE, R. B. (1953), *Scientific Explanation*, Cambridge University Press, Cambridge.

BUNGE, M. (1988), *Ética y Ciencia*, Siglo XX, Buenos Aires.

ECHEVERRÍA, J. (1995), *Filosofía de la Ciencia*, Akal, Madrid.

ECHEVERRÍA, J. (1998*a*), «Ciencia y valores: propuestas para una axionomía de la Ciencia», en MARTÍNEZ FREIRE, P. (ed.), *Filosofía actual de la Ciencia*, Publicaciones de la Universidad de Málaga, Málaga, pp. 175-194.

ECHEVERRÍA, J. (1998*b*), «Teletecnologías, espacios de interacción y valores», *Teorema*, v. 17, n. 3, pp. 11-25.

ECHEVERRÍA, J. (1998*c*), «La emergencia del paradigma postmoderno», en GOMA, J. (1998) (ed.), *Ciencia moderna y postmoderna*, Fundación March, Madrid.

ECHEVERRÍA, J. (2001*a*), «Ciencia, Tecnología y valores. Hacia un análisis axiológico de la actividad tecnocientífica», en IBARRA, A. y LÓPEZ CEREZO, J. A. (eds.), *Desafíos y tensiones actuales en Ciencia, Tecnología y Sociedad*, Biblioteca Nueva/OEI, Madrid, pp. 137-148.

ECHEVERRÍA, J. (2001*b*), «Tecnociencia y sistemas de valores», en LÓPEZ CEREZO, J. A. y SÁNCHEZ RON, J. M. (eds.), *Ciencia, Tecnología, Sociedad y Cultura en el cambio de siglo*, Biblioteca Nueva/OEI, Madrid, pp. 221-231.

GONZÁLEZ, W. J. (1999) (ed.), *Ciencia y valores éticos*, número monográfico de *Arbor*, v. 162, n. 638.

HACKING, I. (1982), *Representing and Intervening*, Cambridge University Press, Cambridge.

HEMPEL, C. G. y OPPENHEIM, P. (1948), «Studies in the Logic of Explanation», *Philosophy of Science*, v. 15, pp. 135-175.

HEMPEL, C. G. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, Nueva York.

HUME, D., *Investigación sobre el conocimiento humano*, Alianza Ed., Madrid, 1980.

KUHN, TH. S. (1977), *The Essential Tension*, University of Chicago Press, Chicago.

LAUDAN, L (1984), *Science and Values*, University of California Press, Berkeley.

LAUDAN, L. (1996), *Beyond Positivism and Relativism*, Westview Press, Boulder.

LÓPEZ CEREZO, J. A. y LUJÁN, J. L. (2001), *Ciencia y Política del riesgo*, Ariel, Barcelona.

MILL, J. S. (1843), *A System of Logic*, John W. Parker, Londres.

NAGEL, E. (1953), «Teleological Explanation and Teleological Systems», en RATNER, S. (ed), *Vision and Action*, Rutgers University Press, New Brunswick (NJ), pp. 192-223.

NAGEL, E. (1977), «Teleology revisited», *Journal of Philosophy*, v. 74, pp. 261-301.

OLIVÉ, L. (2000), *El bien, el mal y la razón*, Paidós, México.

POPPER, K. R. (1935), *Logik der Forschung*, Springer, Viena.

PROCTOR, R. N. (1993), *Value-Free Science? Purity and Power in Modern Knowledge*, Harvard University Press, Cambridge (MA).

PUTNAM, H. (1981), *Reason, Truth and History*, Cambridge University Press, Cambridge.

QUINTANILLA, M. A. (1989), *Tecnología. Un enfoque filosófico*, Fundesco, Madrid.

RESCHER, N. (1987), *Ethical Idealism*, University of California Press, Berkeley.

RESCHER, N. (1988), *Rationality*, Clarendon Press, Oxford.

RESCHER, N. (1993), *A System of Pragmatic Idealism*, vol. II: *The Validity of Values*, Princeton University Press, Princeton.

RESCHER, N. (1999), *Razón y valores en la Era científico-tecnológica*, Paidós, Barcelona.

RUBINSTEIN, A. (1998), *Modeling Bounded Rationality,* MIT Press, Cambridge (MA).

SALMON, W. C. (1990), *Four Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis.

SALMON, W. C. (2002), «Explicación causal frente a no causal», en GONZÁLEZ, W. J. (ed), *Diversidad de la explicación científica*, Ariel, Barcelona, pp. 97-115.

PARTE IV

LAS EXPLICACIONES CAUSALES: ESTRUCTURA Y APLICACIÓN

CAPÍTULO 6

LA ESTRUCTURA DE LA EXPLICACIÓN CAUSAL[1]

WESLEY C. SALMON

En mi trabajo anterior he abordado varios planteamientos no causales de la explicación científica (Salmon, 2002), incluyendo la «Concepción heredada», cuya versión más clara y completa se encuentra en C. G. Hempel, y que fue aceptada por muchos otros. De manera breve, también abordé mi modelo de la relevancia estadística, la concepción unificada de Michael Friedman y la teoría pragmática de [Bas] van Fraassen, señalando serios defectos en todos ellos. Obviamente, hay muchas otras teorías no causales, pero dado el fuerte atractivo intuitivo de la íntima conexión entre causalidad y explicación, analicé brevemente el principal escollo que se interpone en el camino de una teoría causal de la explicación, esto es, el problema de la causalidad suscitado por Hume. Tras haber revisado varias maneras de enfocar ese problema y después de encontrarlas deficientes, deseo ofrecer ahora otro análisis de la causalidad, que considero más aceptable que cualquier otro. Éste es el tema central de esta ponencia.

1. **Un nuevo punto de partida**

La caracterización de la causalidad de Hume incluye un par de sucesos (o hechos) C y E —la causa y el efecto en cuestión— y una relación R que se mantiene entre ellos. La mayoría de los planteamientos posteriores han adoptado el mismo esquema; la cuestión crucial estriba entonces en la naturaleza de esa relación causal R.

1. Ponencia presentada el día 9 de marzo de 2001 en las *Jornadas sobre Variedades de explicación científica*. La traducción ha sido realizada por Wenceslao J. González.

Propongo que, *temporalmente*, dejemos de pensar en hechos o sucesos separados y que, *temporalmente*, eliminemos los términos «causa» y «efecto». En cambio, sugiero que empecemos con la noción de *proceso*. Todos estamos muy familiarizados con numerosos ejemplos de procesos. Por ejemplo, un objeto material en movimiento es un proceso; de hecho, lo es también una partícula que esté en reposo en nuestro marco de referencia particular. Una sombra que se mueve es un proceso. La propagación de una onda es un proceso. Un fotón que se mueve en el espacio es un proceso. Dicho en términos generales, un proceso es algo cuya línea se extiende en el mundo, dentro de un diagrama espacio temporal. Básicamente, los sucesos se representan mediante puntos en diagramas espacio-temporales. Así, los procesos se extienden relativamente en el tiempo y, quizá, también en el espacio. Los sucesos se encuentran localizados relativamente en cosas como un estornudo, una bala que da en su objetivo o el choque entre dos automóviles.

Una pregunta natural acerca de este tema es cómo pueden ser, de grandes o de pequeños, los procesos y los sucesos. La respuesta es netamente pragmática: depende de la naturaleza de nuestras investigaciones. Para un ingeniero que se ocupa del tráfico, por ejemplo, un automóvil en movimiento puede ser un proceso simple. Para un ingeniero de la industria del automóvil, que está interesado en el funcionamiento de los motores de coche, un vehículo en movimiento sería un sistema complejo de procesos e interacciones. Para un astrónomo, un planeta como la Tierra moviéndose alrededor del Sol podría ser un proceso simple. Para un geofísico, podría ser un sistema extremadamente complejo de procesos e interacciones.

En seguida pondré mis cartas sobre la mesa. Deseo mostrar que los *procesos causales* son, precisamente, las conexiones causales que en vano Hume trató de descubrir. En algunos casos, sin embargo, no serán conexiones *necesarias*; cabe admitir también conexiones probabilísticas. Para exponer este asunto, es necesario primero mostrar cómo se puede distinguir entre procesos causales y pseudo-procesos. Ésta es la tarea a la que voy ahora.

2. Procesos causales frente a pseudo-procesos

La faceta más importante de los procesos causales es que *transmiten* algo efectivo (por ejemplo, energía, información, carga eléctrica, momento o influjo causal). Los pseudo-procesos no lo hacen. La cuestión es cómo caracterizar la diferencia. Se han dado dos res-

puestas: una es de Hans Reichenbach (cfr. Reichenbach, 1956 [1928]), y la otra es de Phil Dowe (cfr. Dowe, 1992).

Según Reichenbach, los procesos causales tienen la capacidad de transmitir rasgos (*transmit marks*). Hay un ejemplo simple del tipo que, a menudo, fue usado por Reichenbach y, después, por mí. Considérese un faro que guía los barcos en la noche. Cuando se enciende por la noche, es una guía giratoria, que envía sus haces de luz blanca en direcciones diferentes, según su giro. Cuando el haz de luz encuentra un objeto opaco, como un muro de un edificio o las nubes distantes, cercanas al horizonte, aparece en la superficie un punto (*spot*) de luz blanca que se mueve. La luz que sale hacia fuera del fanal es, claramente, un proceso causal. Como señaló Reichenbach, si pones un filtro rojo en cualquier lugar del haz de luz, el foco de luz que sale del fanal se volverá rojo *y seguirá siendo de color rojo* a partir de ese momento, si no hay ninguna interacción reseñable. Desde luego, puedes hacer que, en algún momento, el punto (*spot*) caiga sobre la superficie opaca roja (p. ej., al colocar allí un trozo de celofán rojo), pero como el punto de luz se mueve, volverá a su color blanco [original], a menos que continúes poniendo marcas, según se mueva de un lugar a otro. La transmisión de rasgos por procesos causales resulta familiar. Una emisora de radio emite una onda que se puede distinguir al modificar su amplitud (AM) o su frecuencia (FM). La transmisión de esos rasgos hace posible, obviamente, el transmitir información. Los procesos fisiológicos se pueden marcar mediante la introducción de trazas radiactivas; esto permite a los fisiólogos saber cómo se dispersan las sustancias en el interior de un organismo.

Mucho más recientemente, Dowe (cfr. Dowe, 1992) ha propuesto una teoría de la transmisión causal de la «cantidad conservada» (*conserved quantity*), que tiene algunas ventajas importántes (para realizar análisis filosóficos) respecto del método de los rasgos (*marks*). Algunas «cantidades conservadas» conocidas son la energía, el momento y la carga eléctrica. La primera ventaja de la teoría de la «cantidad conservada» es que define los procesos causales en términos de características que poseen de modo efectivo (esto es, una o más cantidades conservadas) en lugar de [hacerlo sobre la base de] una mera capacidad (a saber, la *habilidad* de transmitir rasgos). Otras ventajas de la teoría de la «cantidad conservada» aparecerán cuando abordemos las interacciones causales.

Otra diferencia fundamental entre los procesos causales y los pseudo-procesos aparece a partir de la teoría especial de la relatividad. Según esa teoría, ninguna señal puede viajar a una velocidad

más alta que la velocidad de la luz en el vacío. Un mensaje que se transmite por una luz visible o por cualquier otro tipo de radiación electromagnética viaja, obviamente, a la velocidad de la luz. Los objetos materiales, que pueden ser usados para llevar mensajes, pueden ser acelerados hasta llegar a velocidades cercanas a la velocidad de la luz, pero nunca a una velocidad igual o superior a la luz. Los pseudo-procesos, en cambio, pueden viajar arbitrariamente a altas velocidades. Con frecuencia, la gente dice que, según la teoría de la relatividad, nada puede viajar más deprisa que la luz. Esto no es verdad, a menos que, en lo que atañe a nuestro tema, especifiquemos claramente que los pseudo-procesos no son «cosas».

Volvamos al ejemplo del faro. Vamos a imaginarlo rodeado por un sólido muro circular, cuyo radio es igual a la distancia que hay desde el faro hasta la posición de ustedes. El faro proyecta un punto de luz, que recorre el muro en el tiempo que tarda la fuente de luz en hacer una rotación (digamos que una vuelta por segundo). Manteniendo el fanal en el centro del círculo y siendo constante su *ratio* de revolución, imaginemos que ampliamos a un tamaño mayor el muro que lo rodea. Obviamente debe incrementarse la velocidad a la que se desplaza el punto [de luz], puesto que el punto de luz que se mueve a lo largo del muro debe dar una vuelta completa en torno a la circunferencia en el tiempo que tarda el fanal en rotar una vez (un segundo). En la medida en que seguimos ampliando el círculo, llegaremos a un tamaño que requiere que el punto se desplace a una velocidad mayor que la velocidad de la luz.

He aquí otro ejemplo. Un avión que vuela en un día soleado proyecta una sombra sobre el suelo. El avión es un proceso causal; la sombra es un pseudo-proceso. Si dos aviones, volando a diferentes altitudes, pasan simultáneamente por un lugar determinado del terreno, sus sombras se cruzan, pero las sombras pasan por encima del lugar de intersección como si nada hubiera sucedido. Sin embargo, si los aviones hubieran volado a la misma altitud, y hubieran chocado el uno con el otro, el resultado habría sido muy diferente: la destrucción de los aparatos y la muerte de todas las personas que llevaran a bordo.

Para explicar la diferencia entre procesos causales y pseudo-procesos, he hecho uso del concepto de transmisión: la transmisión de rasgos y la transmisión de cantidades conservadas. Puesto que la transmisión es un concepto claramente causal, debemos ahora dedicar nuestros esfuerzos a entender esa noción.

3. La transmisión causal

La distinción entre procesos causales y pseudo-procesos es obviamente fundamental para la teoría de la causalidad que estoy intentando desarrollar, tanto si estamos pensando en términos de rasgos como si lo hacemos a tenor de cantidades conservadas. Volvamos de nuevo al ejemplo del faro. Como los rayos de luz se emiten desde la fuente de luz, la energía está siendo transmitida desde el fanal hacia un muro sobre el que cae la luz. En contraste, el punto de luz que se desplaza a lo largo del muro *tiene* energía, pero no la *transmite*. Ahí está la diferencia básica. Como el haz de luz se emite desde el fanal, la energía se transfiere desde la fuente de luz hasta el haz de luz, pero *no hay una ulterior entrada de energía* cuando el haz de luz se desplaza desde la fuente hasta el muro. El punto que se mueve de una parte a otra del muro tiene también energía, pero ese punto continúa existiendo y posee energía sólo en cuanto que *se le proporciona nueva energía* a cada momento de su trayectoria. En otras palabras, el suministro de energía al punto [de luz] debe ser renovado constantemente, mientras que la energía del haz de luz viaja con el haz sin reaprovisionamiento.

Para aclarar este concepto de transmisión, me gustaría volver atrás cerca de veinticinco siglos a la famosa paradoja de la flecha de Zenón. Según parece, Zenón argumentó que, en cada momento de su trayectoria, la flecha está en reposo. Puesto que, en cada momento, ocupa una cantidad de espacio exactamente igual a su tamaño, no tiene espacio extra en el que moverse. Además, puesto que cada instante es indivisible, no hay tiempo en el que moverse. De ahí que esté siempre quieta, de modo que, posiblemente, no pueda moverse.

Una reacción inmediata moderna ante este argumento consiste en apelar al cálculo infinitesimal y a la distinción que nos permite trazar entre movimiento instantáneo y reposo instantáneo. Sin embargo, Bertrand Russell señaló que esa respuesta a Zenón es realmente una petición de principio (cfr. Russell, 1922). Para encontrar una solución satisfactoria a la paradoja, Russell introdujo lo que denominó la «teoría del movimiento en-en (*at-at*)». Se describe el movimiento estableciendo pares entre diversos puntos espaciales y sus correspondientes momentos de tiempo. Decir que una flecha se mueve desde *A* hasta *B significa* que ocupa cada punto intermedio del espacio *en el* momento apropiado del tiempo. Cabe observar que esta definición de movimiento *no* incluye la no-

ción de velocidad instantánea. No señala que la flecha pase a través de esos puntos con gran rapidez; indica, simplemente, que estaba *en* ese lugar, sin mención alguna a la velocidad. Si se pregunta cómo se desplazó la flecha desde el arquero hasta la mitad de su trayectoria, la respuesta es la misma, a saber: que ocupó los puntos intermedios en los momentos apropiados.

Haciendo uso de la intuición fundamental de Russell, podemos ahora aplicar la «teoría en-en» a la trasmisión causal. Decir que un rasgo se transmite desde *A* (el punto donde se introduce) hasta algún otro lugar, *B*, significa que el rasgo está presente *en* todos los puntos que hay entre *A* y *B*, *sin que se reintroduzca* en ningún sitio del trayecto. Se da *en* el lugar apropiado *en* el tiempo oportuno. Decir que una cantidad conservada, como una carga eléctrica, se transmite desde *A* hasta *B*, en un proceso concreto, significa que la carga eléctrica dada está presente en A y en B, y en todo punto que interviene en el proceso, *sin que sea realimentada* en ninguno de los lugares intermedios.

Para explicar el concepto de transmisión, he hecho referencia a nociones tales como imponer (*imposing*) o volver a imponer (*reimposing*) un rasgo y suministrar o realimentar una cantidad conservada. Esos conceptos son obviamente causales, en la medida en que se incluye una interacción causal a la hora de imponer el rasgo o al transferir la cantidad conservada. Necesitamos ahora dirigir nuestra atención a las interacciones causales.

4. **Interacciones causales**

Al comenzar introduje el concepto de proceso, antes de distinguir entre procesos causales y pseudo-procesos. En ese momento de lo tratado, «proceso» no era una noción causal; de hecho, al hacer referencia a diagramas espacio-temporales, sugerí implícitamente que es una noción geométrica. Siguiendo esa línea de pensamiento, podemos fijarnos en las *intersecciones* de los procesos como una noción geométrica ulterior; son intersecciones de líneas del mundo (*worldline*) en diagramas espacio-temporales. La tarea que encaramos ahora es distinguir entre interacciones *causales* y las meras intersecciones espaciotemporales. A este respecto, necesitamos otro concepto básico, a saber: *cambio*. Es asimismo una noción no causal, en cuanto que los cambios pueden ser causados o no. *Mi cometido es caracterizar las interacciones causales en términos de los conceptos no causales de proceso, intersección y cambio.*

Cuando interseccionan dos procesos, pueden darse cambios en el entorno inmediato de esa intersección. Esos cambios pueden continuar (*persist*) o no una vez pasado el lugar de cruce. Mencioné, por ejemplo, la intersección de las sombras de dos aviones que vuelan a diferente altitud. Aunque las dos sombras cambian de forma en la intersección, continúan después como si no hubiera ocurrido ninguna intersección. Las intersecciones de esa clase no las clasificaremos como interacciones causales. En contraste, cuando chocan las bolas de billar de Hume, los dos procesos cambian, de modo que siguen después del lugar de intersección. Se modifican los momentos de ambas bolas. Una estaba originalmente en reposo; su momento lineal era cero. El momento lineal de la otra bola no era cero; un resultado posible es que todo su momento lineal fuera transferido a la bola que estaba originalmente en reposo, de manera que, la que estaba inicialmente en reposo, ahora se mueve y la otra, originariamente en movimiento, ahora está en reposo.

Estos dos ejemplos nos ofrecen pistas sólidas sobre cómo distinguir entre las interacciones causales y las meras interacciones espacio-temporales. En la interacción causal entre las dos bolas de billar, cada uno de los dos procesos de entrada se modifica de una manera que continúa después del lugar de intersección. Podemos considerar la interacción entre las dos bolas de billar como la imposición (*imposition*) de un rasgo en cada una de ellas, a saber: un cambio en su estado de movimiento. Los rasgos se transmiten, porque la modificación del estado de movimiento persiste sin ninguna influencia externa. Este ejemplo es, en principio, semejante a la coloración de la luz blanca mediante un filtro rojo. En este último caso, el filtro es un proceso y el haz de luz es otro. Cuando intersecan, cambia el color de la luz, y el filtro se torna más cálido al absorber la energía de la luz. Por el contrario, cuando se cruzan las sombras de los aviones, no hay modificación alguna de la sombra que siga después del lugar de intersección. Esta intersección *no* es una interacción causal.

Otra forma de describir esos ejemplos es en términos de la transferencia de cantidades conservadas. Cuando chocan las bolas de billar, hay una transferencia del momento lineal de la una a la otra. Cuando la luz pasa a través del filtro, hay una transferencia de energía desde la luz al filtro. Si pensamos en los procesos causales en términos de su capacidad para transmitir rasgos, podemos decir que una *interacción causal* es una intersección en la que cada uno de los procesos entrantes queda marcado (esto es, modificado), y el rasgo continúa después de la intersección sin [la presencia de] in-

teracciones posteriores que marquen. Si pensamos en los procesos causales en términos de transmisión de cantidades conservadas, podemos decir que una *interacción causal* es una intersección de procesos donde los procesos de salida cuentan con una suma de cantidad conservada que difiere de lo poseído por los procesos de entrada. En contraste, las sombras de los aviones no poseen cantidades conservadas; las cantidades que de hecho muestran, como la forma y el tamaño, no se conservan. *A fortiori*, cuando las sombras interseccionan, no hay intercambio de cantidades conservadas.

En este punto, podemos ver otra ventaja importante del enfoque de la cantidad conservada sobre el método del rasgo (*mark method*). En las intersecciones consideradas hasta aquí, teníamos dos procesos de entrada y dos procesos de salida. Vamos a llamarlos «intersecciones de tipo-X». Por supuesto, hay otras clases de intersecciones. Por ejemplo, hay casos en los que un proceso se divide en dos, sin que incida ningún otro proceso. Cuando una gallina pone un huevo, hay inicialmente un solo animal; posteriormente, hay un animal y un huevo. En este caso no relativista (*Nonrelativistic*), podemos considerar la masa como la cantidad conservada. La masa de la gallina justo antes de poner el huevo no es igual a su masa después de hacerlo o a la masa del huevo. Desde luego que la masa de la gallina más la masa del huevo justo después de la puesta es igual a la masa de la gallina justo antes de la puesta, pero esto es, simplemente, reformular la conservación de la masa. Vamos a llamar «intersecciones tipo-Y» a las intersecciones de esta clase, pensando en un diagrama espacio-temporal en el que el tiempo discurre desde la parte inferior a la parte superior (como es costumbre). Otra intersección del tipo-Y sucede cuando un átomo en un estado acelerado (*excited*) emite un fotón, provocando así una transición a un estado de menor energía. Al principio tenemos un proceso (el átomo en el estado de aceleración), que entonces se divide en dos (un fotón y un átomo con menos energía). En este caso, la energía es la innegable cantidad conservada.

Los átomos también absorben fotones, transformándolos en estados de más energía. En este caso, tenemos dos procesos que se fusionan y salen como uno. De nuevo, la energía es la cantidad conservada relevante. Podemos denominar «intersecciones tipo-L» a las intersecciones de esta clase. Las consideraciones se reflejan en aquellas intersecciones de tipo-Y. Cuando una serpiente engulle un ratón, tenemos otro ejemplo de la intersección tipo-L: presenta una fuerte analogía [con el caso de] la gallina poniendo un huevo. La masa es, de nuevo, la cantidad conservada.

Me parece que la teoría de la cantidad conservada maneja las intersecciones de los tipos Y y L de una manera que va directamente al grano: cuando se fusionan (*merge*) dos procesos o cuando surgen de uno, la cantidad conservada en cuestión es diferente en los procesos de entrada o de salida. De este modo, se intercambia una cantidad conservada. No veo cómo podría el método del rasgo tratarlos de manera adecuada. Así, en lo que atañe al análisis filosófico, ahora defiendo la teoría de la cantidad conservada (*conserved-quantity*) de Dowe, pero sigo conservando el método del rasgo (*mark method*) como una herramienta eficaz para descubrir y verificar (*veryfing*) relaciones causales. Las interacciones X, Y y L son los tipos básicos; obviamente, existen tipos más complicados. No veo razón por la que no haya seis procesos de entrada y cuatro de salida, sólo por tirar los números un poco hacia arriba.

5. **Causas y efectos**

Al principio del tema sugerí que abandonáramos *temporalmente* los términos «causa» y «efecto». Ha llegado el momento de *considerar* su reintroducción. Como veremos, aunque pueda parecer simple el volver a traerlos a colación, cuesta un poco de trabajo.

Me parece que hay, al menos, dos modelos (*patterns*) básicos que, de modo natural, podemos calificar como relaciones *directas* de causa-efecto. El primero y más simple es el choque entre las bolas de billar de Hume. Es sólo una interacción causal entre dos procesos. La interacción produce cambios en los dos procesos, pero, puesto que estamos más interesados en el cambio desde el reposo al movimiento en lo que atañe a la segunda bola, decimos que el choque *causó* el movimiento de la segunda bola. En un contexto poco diferente, podríamos estar más interesados en el posterior movimiento de la primera bola. En una interacción del tipo-X se modifican los dos procesos de entrada, y estas dos modificaciones son el efecto de la interacción. El contexto determina qué suceso (*event*) tomamos como efecto. Para los jugadores experimentados, el giro de la bola blanca es un asunto serio, porque determina en gran medida lo que le ocurre a la bola blanca tras la colisión. Esto significa que, además del momento lineal, ha de conservarse el momento angular. Así, el grado de pericia de los jugadores influye también en los aspectos a los que ellos atienden. Es evidente que al ejemplo clásico de Hume lo rodea un alto grado de dependencia del contexto.

El segundo tipo es un poco más complejo. Este modelo de causa-efecto consiste en una *interacción causal* entre dos procesos, un *proceso causal* que surge a partir de la interacción, y una *interacción* posterior de ese proceso con otro proceso. Las interacciones se pueden considerar sucesos, de modo que este modelo consta de dos sucesos conectados por un proceso causal. Por ejemplo, un chaval golpea con un bate una bola de béisbol que le lanzan, la bola cambia su dirección y vuela hacia una ventana, y entonces la bola rompe la ventana haciéndola añicos. En palabras de Hume: «éste es un ejemplo tan perfecto del (segundo tipo de) relación de causa y efecto como alguno que conocemos por sensación o por reflexión». Hablando coloquialmente, el muchacho que golpea la bola con el bate *causa* la rotura de la ventana (el *efecto*). En este caso, el bate y la bola son dos procesos separados que interseccionan e interactúan causalmente; la bola es un proceso causal que se mueve desde el bate a la ventana; y el proceso causal constituido por la bola de béisbol que se desplaza intersecciona e interactúa causalmente con el proceso constituido por el cristal de la ventana. La bola de béisbol que se mueve *es* la conexión causal entre el golpe de la bola y la rotura de la ventana. Este proceso causal *transmite* masa, momento y energía de un suceso al otro; así es cómo un proceso causal *propaga* influencia causal. Este ejemplo muestra dos características de la causalidad: *producción* y *propagación*. Las interacciones causales *producen* cambios en procesos causales; los procesos causales *propagan* los resultados.

Dados dos sucesos, separados espacio-temporalmente el uno del otro, ¿bajo qué circunstancias podemos decir que están conectados *directamente* como causa y efecto? Es bastante fácil encontrar algunas condiciones necesarias. Una condición necesaria es que haya un proceso causal que se extienda del uno al otro. Al mismo tiempo, es evidente que esto no puede ser una condición suficiente. Recordemos el ejemplo del béisbol. Supongamos que, justo cuando la bola va a romper la ventana, el bateador se anticipa a lo que está a punto de ocurrir y grita: «¡oh, no!» (o algo parecido), de modo que las ondas sonoras de la exclamación alcanzan la ventana precisamente cuando lo hace la bola. Reconocemos inmediatamente que, tanto la onda sonora como la bola, portan un momento lineal, pero el momento de la onda sonora es insuficiente para romper la ventana. Una clara limitación (*constraint*) aquí es la conservación del momento lineal. Cuando pensamos que se puede incumplir (*might be violated*) la condición de conservación, dejamos a un lado el juicio según el cual se logra una relación causa-efecto. Parece que algo

esencial ha quedado fuera de lo narrado. No incumplir (*nonviolation*) las relaciones de conservación es otra condición necesaria.

Por poner otro ejemplo: consideremos el granero que se incendió. Tenemos aquí una estructura más complicada, que incluye procesos e interacciones. Un cigarrillo encendido se cae de la mano de un vagabundo y va a dar sobre algo de paja seca del suelo. En la interacción entre esos dos procesos se intercambia energía, subiendo la temperatura de la paja hasta su temperatura de ignición. El trozo de paja que se quema libera energía que enciende los trozos de paja vecinos. Se sigue una conflagración, en la que la paja que arde quema trozos cercanos de madera; de esta manera, el granero entero se consume en llamas. La serie de procesos e interacciones de este ejemplo se comprende bien científicamente; está claro cómo se transmite la energía mediante procesos y cómo se intercambia a través de interacciones. El suceso completo es extremadamente complejo, pero no hay nada particularmente problemático en él. La caída del cigarrillo encendido causó el incendio del granero.

Este ejemplo ilustra los conceptos de Mackie de la condición INIS y la *causa plena* (o «filosófica»).[2] Recuerden que una condición INIS es una causa en su acepción ordinaria. En este caso, el lanzamiento del cigarrillo encendido es la causa de la quema del granero. No hay duda que la selección de una condición INIS u otra es netamente pragmática y que depende del contexto. En algunos casos, tales como una chispa de la antorcha de un trabajador, se señalaría la presencia de material combustible como la causa. La causa plena es una disyunción que engloba elementos tales como el lanzamiento de un cigarrillo encendido, una chispa de la antorcha de un trabajador, ser alcanzado por un rayo, la combustión espontánea del heno almacenado en el granero, y un acto deliberado de un pirómano. Cabe recordar que cada uno de los disyuntos contiene otros elementos suficientes para constituir una condición suficiente para el efecto. Esta compleja fórmula enuncia una condición necesaria y suficiente para el efecto, esto es, la quema del granero.

El enfoque de Mackie requiere dos comentarios. Primero, que en muchos casos, y éste no es una excepción, no se dan conjuntos completos de disyuntos. Por ejemplo, no se ha mencionado el material incendiario caído en el tejado desde un cercano incendio forestal. Mackie admite explícitamente que las formulaciones de las causas plenas

2. Como se expuso en el trabajo anterior, la condición INIS es «una parte **I**nsuficiente, **N**o redundante (esto es, necesaria), de una condición **I**nnecesaria y **S**uficiente»; cfr. SALMON, W. C., «Explicación causal frente a no causal», pp. 109-112. *(N. del t.)*

contienen habitualmente «proposiciones universales *elípticas* o *con huecos*» (cfr. Mackie, 1974, p. 66); los huecos son el reflejo de las condiciones suficientes que, en un momento determinado, son desconocidas. Es difícil ver cómo esas proposiciones pueden constituir descripciones objetivas de causas plenas; se encuentran *relativizadas epistémicamente* respecto de las situaciones de conocimiento de las personas implicadas. Segundo, Mackie mantiene explícitamente que el enunciado de la causa plena formula condiciones tanto necesarias como suficientes para el efecto *ante el trasfondo de un campo causal.* Entre otras cosas, el campo podría cubrir condiciones como la presencia de oxígeno. De nuevo, hay circunstancias especiales (un experimento particular de laboratorio, por ejemplo) en las que la presencia de oxígeno sería singularizada como la causa. Parece obvio que la selección del campo causal se guía por consideraciones pragmáticas y, por tanto, depende del contexto.

A la vista de los ejemplos anteriores, como también de muchos otros, concluyo que los enunciados de causa-efecto son casi siempre (si no siempre) dependientes del contexto. Si esta conclusión es correcta, significa que la mayoría de los enunciados causa-efecto carecen de objetividad completa (*full objectivity*). Dada esta conclusión, parece que afrontamos un serio problema al intentar proporcionar un enfoque de la causalidad completamente objetivo. Sin embargo, pienso que puede hacerse, pero sólo yendo a un nivel diferente.

6. **Estructura causal completa**

El concepto de una relación directa causa-efecto comporta mucha finura. Para facilitar una ulterior aclaración, introduciré el concepto de una *estructura causal completa*. Podemos hacerlo con bastante facilidad al tener a nuestra disposición los conceptos de *proceso causal, transmisión causal* e *interacción causal.* La estructura causal completa de cualquier región espacio-temporal (esto es, del universo) se da mediante la red completa de procesos causales y de interacciones causales contenidas en esa región elegida. Debe incluir un recuento de las cantidades conservadas transmitidas por los procesos y de las intercambiadas en las interacciones. Habrá que dar cuenta de los procesos que entran o salen de esa sección, y de las cantidades conservadas que ponen o sacan.

Mi noción de estructura causal completa está estrechamente relacionada con el concepto de *texto explicativo ideal* de Peter Railton (cfr. Railton, 1981), que podría incluir todos los hechos —no im-

porta qué insignificantes puedan ser— que, en algún sentido, sean relevantes para el *explanandum* en consideración. En el caso de la explicación causal, el texto explicativo ideal sería esencialmente el mismo que la estructura causal completa. La diferencia fundamental entre ellos es que la construcción de Railton es un *texto* (una entidad lingüística), mientras que mi estructura causal completa es una *entidad física* compleja. Railton señaló con énfasis que rara vez, si es que alguna, intentamos escribir (*write out*) el texto ideal; buscamos más bien el iluminar partes concretas respecto de las cuales de hecho estamos interesados. La selección de la parte a iluminar es claramente una cuestión pragmática, y distintos investigadores elegirán diferentes aspectos para examinar. Para llevarlo a cabo, intentamos aducir *información explicativa*, lo que comporta seleccionar alguna parte o aspecto del texto ideal.

Una característica básica compartida por el texto explicativo ideal de Railton y mi estructura causal completa es su objetividad. Son correctas o incorrectas sin tener en cuenta el contexto u otras consideraciones pragmáticas. Éste es un aspecto fundamental de mi propuesta para articular un enfoque objetivo de la causalidad. La estructura causal completa es un hecho de la Naturaleza que existe de manera totalmente independiente respecto de nuestro conocimiento o nuestros intereses; no está relativizado epistémicamente. Los enunciados sobre las relaciones entre causas y efectos son, por lo general, altamente selectivos, y habitualmente dependen de un contexto. Por ejemplo, cuando tratamos la rotura de una ventana por una bola de béisbol, pasamos por alto los múltiples choques de la bola con las moléculas de aire; es suficiente considerar el momento de la bola después de haber interactuado con el bate. Los choques con las moléculas son parte de la estructura causal completa, pero no son pertinentes para la exposición. Sin embargo, si deseamos tratar de la curva de la bola lanzada por el *pitcher*, [entonces] las interacciones entre la bola y el aire son esenciales.

Desde el principio, Mackie estaba profundamente interesado en encontrar la causalidad «en los objetos» (esto es, en proporcionar un planteamiento de la causalidad completamente objetivo, no dependiente de la mente; cfr. Mackie, 1974, pp. 1-2). Este enfoque tiene dos partes: 1) la especificación de la prioridad causal sin [hacer] referencia a la prioridad temporal; y 2) el concepto de una condición INIS. En mi primera ponencia,[3] en relación con el ejemplo de la máquina de bolas (*pinball machine*), mostré que su análisis de la

3. Cfr. SALMON, W. C., «Explicación causal frente a no causal», pp. 111-112. *(N. del t.)*

prioridad causal en términos de fijación causal es totalmente insostenible. No obstante, las bifurcaciones conjuntivas (*conjunctive forks*) de Reichenbach podrían ser traídas para determinar la prioridad causal.

Las condiciones INIS, a mi juicio, proporcionan una herramienta útil para el análisis de las relaciones causa-efecto, pero Mackie insiste correctamente en que no son completamente objetivas ni independientes del contexto. Las «proposiciones universales *elípticas* o con huecos (*gappy*)» de Mackie, que se diseñaron para formular la *causa plena* objetiva, todavía dan una fuerte apariencia de relatividad epistémica o de dependencia del contexto (cfr. Mackie, 1974, p. 66). En las páginas posteriores, da buena cuenta de la utilidad práctica de las condiciones INIS; esto muestra que tienen virtudes pragmáticas significativas, pero no da cuenta de manera convincente de la completa objetividad de su causa plena.

A lo largo de todo *El cemento del Universo*, Mackie hace múltiples referencias casuales, más o menos de pasada, a los procesos causales, los mecanismos causales, las interacciones y las configuraciones con causa común. Lo que no consigue ver, en mi opinión, es el hecho de que ésas son, precisamente, la clase de entidades que forman la base objetiva de nuestras propuestas causales. Además, manifiestan otro de los principales *desiderata* de Mackie, a saber: la capacidad para distinguir entre secuencias de sucesos (*events*) causales y no causales (cfr. Mackie, 1974, p. 29). En un punto anterior, me he ocupado de mostrar cómo podemos hacer distinciones objetivas entre procesos causales y pseudo-procesos y entre interacciones causales y meras intersecciones espacio-temporales. Mi conclusión es que Mackie llevó a cabo su análisis en un nivel relativamente superficial, en donde las consideraciones pragmáticas y la dependencia del contexto juegan un papel legítimo, pero ignoró casi por completo la base objetiva y no contextual que subyace a su nivel de análisis.

Tal como veo el asunto, la estructura causal completa se limita a las situaciones que no implican fenómenos cuánticos, donde —como es bien sabido— las exposiciones causales encuentran serias dificultades. Debajo de la estructura causal completa hay, por así decir, una cierta subestructura de Mecánica Cuántica. No afirmo que comprenda esos mecanismos cuánticos, más allá de estar profundamente convencido de que no conforman las especificidades de la causalidad normal.

7. La objetividad del enfoque

Para hacer explícito el carácter objetivo del enfoque de la causalidad que he desarrollado, recapitularé la exposición en orden más o menos inverso a como la he presentado. El orden precedente era heurístico; la recapitulación es constructiva. Comenzamos con los *conceptos no causales* de *proceso* e *intersección*. Desde el punto de vista de los diagramas espacio-temporales, son conceptos esencialmente geométricos. En el curso de la exposición, he utilizado conceptos netamente empíricos, tales como propiedades y cambios. He introducido las cantidades como propiedades de objetos o de procesos; algunas cantidades se conservan de hecho (*as a matter of fact*).

Estamos preparados ahora para introducir una secuencia de conceptos *causales*. El primero de ellos es *interacción causal*. Es una intersección espaciotemporal de procesos. Una intersección de este tipo cumple los requisitos de la interacción causal si, y sólo si, los procesos que interseccionan llevan a cabo intercambios de cantidades conservadas en el lugar de intersección, [intercambios] que —en ausencia de intersecciones adicionales— continúan durante algún tiempo más lejos de ese lugar. La persistencia no necesita ser muy duradera; puede ser bastante breve, porque las interacciones causales adicionales pueden ocurrir con frecuencia.

El segundo concepto causal que se introduce es *transmisión causal*. Una cantidad conservada se transmite entre los puntos *A* y *B* en un proceso si, y sólo si, está presente *en la misma medida* en todos los puntos entre *A* y *B*, sin [que haya] una interacción adicional entre esos puntos. De modo semejante, un cambio (o un rasgo) se transmite entre los puntos *A* y *B* si, y sólo si, está presente en todos los puntos entre *A* y *B*, sin una interacción adicional entre esos puntos.

El tercer concepto causal es *proceso causal* (como opuesto a pseudo-proceso). Un proceso es causal si, y sólo si, transmite *realmente* una o más cantidades conservadas. Una cantidad conservada es una propiedad de todo aquello que la posea. La *propagación* causal se identifica con la transmisión por medio de los procesos causales.

El cuarto concepto causal es *conexión causal*. Existe una conexión causal entre dos sucesos si, y sólo si, hay entre ellos un proceso causal, o un conjunto de procesos causales e interacciones, tales que todas las cantidades conservadas se conservan *sin tener en cuenta una interacción adicional*. Como dije antes, un fallo de conservación frustraría cualquier propuesta de identificación de conexiones causales. Considero que una conexión causal, en el sentido

especificado, cumple el *desideratum* de Hume de una conexión entre causa y efecto.

Un quinto concepto causal es aquel de una *relación directa causa-efecto*. He sugerido que las relaciones causa-efecto son, al menos, de dos clases: interacciones causales (por ejemplo, los choques de las bolas de billar) o pares de sucesos (por ejemplo, un bateador que golpea una bola de béisbol que rompe una ventana) conectados por conexiones causales. Ambos tipos cumplen los requisitos de las relaciones directas causa-efecto. En lo que atañe a este estudio, estimo que los sucesos producidos por una causa común no tienen una relación *directa* causa-efecto entre ellos: su relación causal es indirecta.

Hay una cautela *importante*. En modo alguno he pretendido proporcionar un análisis conceptual de la causalidad que pudiera aplicarse a todos los mundos posibles. Más aún, no pretendo sugerir que la causalidad, como ha sido caracterizada aquí, se aplique a todos los dominios de nuestro mundo real. No creo que, lo que cabría llamar «causalidad normal» (*normal causation*), se aplique en el dominio de la Mecánica Cuántica (el dominio —cabría indicar— donde la dualidad onda/partícula se manifiesta en cuanto tal). Me doy cuenta de la presencia de un aspecto ondulatorio en las bolas de billar, pero cuestiono que esto se manifieste de una manera que sea empíricamente significativa.

Reconozco que la teoría que estoy proponiendo tiene un tono altamente reduccionista. Me parece que el enfoque podría sostenerse en las Ciencias de la Naturaleza, incluida la Biología, pero no en la Mecánica Cuántica. No tengo confianza en que sea apropiado para la Psicología o las Ciencias Sociales. Tengo serias reservas con respecto a los fenómenos psicológicos. En lo que atañe a las relaciones interpersonales, sólo mantendría que los procesos causales, la transmisión y la interacción son necesarios para el reconocimiento de los otros y para la comunicación entre ellos. Dejo como cuestión abierta si hay otros tipos de causalidad implicados.

8. **Explicación causal**

En mi libro de 1984[4] distinguí tres visiones generales de la explicación científica: la *epistémica*, la *modal* y la *óntica*. La «Concep-

4. Cfr. SALMON, W. C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton (N. J.). *(N. del t.)*

ción heredada» de Hempel es la posición más representativa de la postura epistémica. Según el planteamiento modal (que he preferido no abordar en estas ponencias), explicar significa mostrar que lo que *de hecho* sucedió *tenía que* suceder a la vista de los hechos explicativos. He apoyado la postura óntica, que, a diferencia de la posición epistémica, considera que las explicaciones son completamente objetivas (esto es, no son relativas a nuestra situación de conocimiento). Desde luego que depende de nuestro estado de conocimiento lo que nosotros *consideramos que son* explicaciones correctas.

En estas dos ponencias, he argumentado a favor de una postura óntica de la causalidad. Al analizar la *estructura causal completa* en términos de interacciones (causales), transmisiones (causales) y procesos causales, he intentado extender el fundamento (*groundwork*) objetivo de la causalidad. Al mismo tiempo, he sostenido que las afirmaciones de causa-efecto dependen habitualmente (si no siempre) de consideraciones contextuales. Así como es un error mantener que hemos encontrado *la única* explicación correcta de un hecho dado, es también un error defender que hemos encontrado *la única* causa de un efecto dado. Sin embargo, *una vez que se ha proporcionado un contexto* mediante las diversas características pragmáticas de la situación, podemos determinar, a través de la referencia a la estructura causal completa, si una enunciación causal es correcta o incorrecta. De modo semejante, *una vez que se ha dado un contexto*, podemos determinar si es correcta o incorrecta una explicación particular.

Dentro del ámbito de las dos ponencias, no es posible entrar en detalle en los factores pragmáticos que determinan los contextos; pero permítanme terminar mencionando brevemente un ejemplo. Ocurre un accidente de avión; pueden estar presentes muchos factores causales. Tal vez hubo una rotura mecánica de una pieza. Quizá las condiciones meteorológicas jugaron un papel. Cabe que se produjera un error del piloto. Quizá una radiación electromagnética externa interfirió con los instrumentos de vuelo. Y así, sucesivamente. Más de uno de estos factores podría haber tenido parte en el accidente (por ejemplo, puede ser que el piloto debiera haber sabido cómo afrontar el tiempo adverso o el fallo de la pieza). La gente que diseña aviones podría centrarse en un defecto de una pieza. Las personas que entrenan a los pilotos podrían concentrarse en los errores del piloto. Los controladores de vuelo prestarían atención a las condiciones meteorológicas. Los aseguradores tratarían de especificar las responsabilidades legales. Y así sucesivamente.

En suma, estoy de acuerdo con Scriven cuando mantiene que identificar una causa es, en muchas situaciones, equivalente a dar una explicación. Mi cometido principal ha sido proporcionar un enfoque de la causalidad —por completo ajeno a su teoría— que haga justicia a los aspectos objetivos y de dependencia del contexto de la causalidad y de la explicación, y también a la crítica clásica de Hume a la causalidad.

Referencias bibliográficas

DOWE, PH. (1992), «Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory», *Philosophy of Science*, v. 59, pp. 195-216.

DOWE, PH. (2000), *Physical Causation*, Cambridge University Press, Cambridge.

FRAASSEN, B. C. VAN (1980), *The Scientific Image*, Clarendon Press, Oxford.

FRIEDMAN, M. (1974), «Explanation and Scientific Understanding», *Journal of Philosophy*, v. 71, pp. 5-19.

HEMPEL, C. G. (1965), *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, Free Press, Nueva York.

HUME, D. (1888 [1739-40]), *A Treatise of Human Nature*, edición de L. A. Selby-Bigge, Clarendon Press, Oxford.

HUME, D. (1955 [1740]), «Abstract of *A Treatise of Human Nature*», en edición de C. W. Hendel (1955).

HUME, D. (1955 [1748]), *An Inquiry Concerning Human Understanding*, edición de C. W. Hendel, Bobbs-Merrill, Indianápolis. [La grafía moderna de «Enquiry» ha sido introducida por el editor.]

HUME, D. (1955 [1776]), «My Own Life», edición de C. W. Hendel, Bobbs-Merrill, Indianápolis.

KANT, I. (1929 [1781]), *The Critique of Pure Reason*, traducción de N. K. Smith, Macmillan, Londres.

MACKIE, J. L. (1974), *The Cement of the Universe*, Clarendon Press, Oxford.

RAILTON, P. (1981), «Probability, Explanation, and Information», *Synthese*, v. 48, pp. 233-256.

REICHENBACH, H. (1958), *The Philosophy of Space and Time*, Dover, Nueva York. Edición del original en alemán: 1928.

REICHENBACH, H. (1956), *The Direction of Time*, University of California Press, Berkeley y Los Ángeles.

RUSSELL, B. (1922), «The Problem of Infinity Considered Historically», en RUSSELL, B., *Our Knowledge of the External World*, George Allen and Unwin, Londres, pp. 159-188.

SALMON, W. C. (1981), «Causality: Production and Propagation», en ASQUITH, P. D. y GIERE, R. N. (eds.), *PSA 1980*, Philosophy of Science Association, East Lansing, Michigan, pp. 49-69. [Reimpreso en Salmon, 1998.]

SALMON, W. C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton (N. J.).

SALMON, W. C. (1994), «Causality Without Counterfactuals», *Philosophy of Science*, v. 61, pp. 297-312.

SALMON, W. C. (1998), *Causality and Explanation*, Oxford University Press, Nueva York.

SALMON, W. C. (1997), «Causality and Explanation: A Reply to Two Critiques», *Philosophy of Science*, v. 64, pp. 461-477.

SALMON, W. C. (2002), «Explicación causal frente a no causal», en GONZÁLEZ, W. J. (ed.), *Diversidad de la explicación científica*, Ariel, Barcelona, pp. 97-115.

SCRIVEN, M. (1958), «Definitions, Explanations, and Theories», en FEIGL, H. ET AL. (eds.), *Minnesota Studies in the Philosophy of Science II*, University of Minnesota Press, Minneapolis, pp. 99-195.

SCRIVEN, M. (1959), «Explanation and Prediction in Evolutionary Theory», *Science*, v. 30, pp. 477-482.

CAPÍTULO 7

LA EXPLICACIÓN CAUSAL EN LAS CIENCIAS SOCIALES[1]

MERRILEE H. SALMON

1. Introducción

Uno de los problemas más engorrosos (*vexing*) de la Filosofía de las Ciencias Sociales es si son posibles las explicaciones causales de la conducta humana. El problema está conectado con profundos desacuerdos filosóficos acerca de las teorías generales de la causalidad y la explicación, del holismo y el individualismo, del libre albedrío y la responsabilidad, y, especialmente, de la causalidad psicofísica. Algunos teóricos de la regularidad causal, como John Stuart Mill (cfr. Mill, 1874), consideran que es un problema abordable. La observación empírica y los estudios históricos apoyan ciertas conexiones regulares entre la información sobre las razones individuales (creencias y deseos) y la conducta. Mill piensa que esas regularidades se pueden aclarar e inferir a partir de regularidades psicológicas más básicas —establecidas asimismo empíricamente—, que conectan los tipos de personalidad con pautas específicas de la conducta. Según J. S. Mill, tras las leyes de naturaleza personal (*character*) se encuentran regularidades psicológicas fundamentales, tales como la ley de asociación, que rigen cómo la mente humana organiza la experiencia. Mill admite que somos incapaces de hacer predicciones precisas de la conducta humana, pero lo atribuye a la multiplicidad de fuerzas causales que interactúan y se contrarrestan, junto con condiciones iniciales cambiantes o no bien conocidas; problemas que afectan, asimismo, a algunas Ciencias Físicas,

1. Ponencia presentada el día 9 de marzo de 2001 en las *Jornadas sobre Variedades de explicación científica*. La traducción ha sido realizada por Wenceslao J. González.

como la Meteorología. Así, Mill argumenta en favor de un modelo de cobertura legal (*covering-law*) para la explicación de la conducta humana, donde las leyes explicativas son causales, aunque, en su mayor parte, sólo proporcionan predicciones imprecisas.

Emile Durkheim (cfr. Durkheim, 1897, 1951), aunque discrepa del psicologismo de J. S. Mill, tampoco ve impedimento conceptual alguno para estipular leyes de la conducta humana. Durkheim, que dominaba mejor que Mill el razonamiento estadístico, mantuvo que los hechos sociales, tales como la *proporción* de suicidios en una sociedad dada, no se puede explicar del todo en términos de las creencias y los deseos de los individuos. Esos fenómenos necesitan causas sociales para su explicación. Por causas sociales, Durkheim parece entender aquellos mecanismos que operan en el ámbito social más que en el ámbito individual. Usó los datos estadísticos para tratar de descubrir mecanismos que subyacen en las proporciones (*rates*) variables de suicidios, tales como los diversos grados de integración social en sociedades que tienen confesiones religiosas diferentes. Aun cuando es cierto que Durkheim consideró esos mecanismos como importantes, parece que la fuerza explicativa de las leyes no depende del todo de esa vía de comprensión (*understanding*) [del fenómeno]. Ante la ausencia de mecanismos plausibles, Durkheim hizo referencia a un «impulso colectivo» que no era meramente un resumen de las diversas intenciones individuales. Por ejemplo, aun cuando no haya mecanismos simples que hagan inteligible la regularidad de la proporción anual de las «cartas de suicidio» en Francia, esta regularidad estadística permite sorprendentemente una predicción precisa del fenómeno social (y tal vez incluso una forma débil de explicación).

Desde Dilthey en adelante, la oposición interpretacionista (*interpretivist*) a la posibilidad de explicación causal de la conducta humana insiste en que la relación entre las razones (esto es, creencias, deseos, intenciones) y la conducta no es causal, al menos no en el mismo sentido de «causal» que opera en el mundo físico. Algunos interpretacionistas, como R. G. Collingwood (cfr. Collingwood, 1946) y P. Winch (cfr. Winch, 1958), mantienen esta posición porque creen que las razones están conectadas *lógicamente* con la conducta, le dan sentido y la hacen ser la clase de conducta que es. Siguiendo a D. Hume, argumentan que esas «cuestiones de razón» no pueden ser «cuestiones de hecho». Así, sostienen que, a menos que «causa» se use en un sentido especial, supone un error categorial entender las razones como causas de la conducta humana. Otros interpretacionistas «más suaves», como el antropólogo C. Geertz (cfr.

Geertz, 1973), rechazan la explicación causal no sobre fundamentos lógicos, sino a tenor de la complejidad y contumacia de la naturaleza humana, la capacidad humana de elección libre o el presunto fracaso para encontrar conexiones entre las razones y la conducta en términos de leyes que sean no triviales.

Tanto los naturalistas, que apoyan la explicación causal de la conducta humana, como sus oponentes se complacen con el argumento de D. Davidson (cfr. Davidson, 1980), según el cual los interpretacionistas como Winch no han logrado deslindar las relaciones causales en el mundo respecto de las descripciones de esas relaciones, algunas de las cuales están estructuradas en oraciones analíticas. Davidson cree que, aunque las razones causan la conducta, puede no haber leyes psicofísicas, porque las descripciones de las razones ni se expresan ni se pueden traducir a una descripción de los fenómenos físicos, que son las causas genuinas de la conducta. No abordaré aquí los méritos del argumento de Davidson o sus críticas (véase, no obstante, McIntyre, 1998), porque deseo, en su lugar, evitar ese asunto para examinar las explicaciones de la conducta humana que no apelan a razones.

Los dos —los interpretacionistas y sus oponentes— están de acuerdo en que el punto central de las Ciencias Sociales es la *acción* humana, a la que distinguen de los meros procesos físicos del cuerpo humano. Entre estos últimos se encuentran los procesos digestivos, la respiración normal y otros procesos del estilo que son más o menos automáticos y que, normalmente, no están bajo control consciente. En muchos casos, está clara la diferencia entre las acciones y los meros procesos físicos del cuerpo. Contener deliberadamente la propia respiración para escapar de una detención es una acción, mientras que la respiración normal es meramente un proceso físico. Mill, un naturalista, se ocupó ciertamente de explicar la conducta que resultaba de las creencias y los deseos de los agentes, no de los procesos fisiológicos. Collingwood, un interpretacionista, caracterizaba las acciones humanas dotadas de algo interno (*inside*) además de lo externo (*outside*). Mediante esta frase, quiere señalar que los eventos no pueden ser descritos de modo completo sin referencia al pensamiento del agente. Mantuvo que el historiador —su término para los que practican cualquier rama de las Ciencias Sociales— se ocupa «de aquellos sucesos que son la expresión externa de su pensamiento, y sólo se ocupan de éstos en la medida en que expresan pensamientos». Señala que «el historiador no está interesado en el hecho de que los hombres (*men*) coman, duerman, hagan el amor y satisfagan de este modo sus apetitos naturales, sino

que se interesa por las costumbres sociales que crean mediante su pensamiento como marco teórico dentro del cual esos apetitos encuentran satisfacción de una manera estipulada mediante la convención y la moralidad» (Collingwood, 1946, p. 216).

Esas caracterizaciones de la acción humana de Mill y Collingwood resultan en cierto modo poco claras en lo que respecta a los límites de la conducta implicada. Por ejemplo, ¿al historiador le atañe sólo la conducta en la que el agente tiene conciencia refleja de lo que intenta cuando realiza la acción? Aunque Mill y otros parecen centrarse en esas acciones, éstas constituyen de hecho una proporción relativamente pequeña de lo que normalmente llamamos acciones humanas. La conciencia refleja (*conscious awareness*) —o prestar atención a lo que estamos haciendo— requiere un esfuerzo considerable y, muy a menudo, se acompaña de intentos de resolver un problema nuevo o de tomar una decisión importante que requiere procesar nueva información. Aprender nuevas destrezas, tales como conducir, requiere mucha atención consciente, pero una vez que se domina esa habilidad, su ejercicio llega a ser automático. Gran parte de nuestra vida social es, de modo similar, automática. Poco se conseguiría si tuviéramos que prestar atención consciente a toda nuestra conducta social. Quizá podríamos decir que las acciones son conductas que, en algún momento, requirieron atención consciente, tales como conducir, aunque más tarde puedan llegar a ser automáticas. Sin embargo, incluso es difícil llegar a plantear que toda nuestra conducta social necesitó atención consciente en la fase de aprendizaje. Muchos psicólogos han argumentado que buena parte del aprendizaje se produce en ausencia de atención consciente (cfr. Hefferline et al., 1959, citado en Gaulin y McBurney, 2001, p. 114). Adquirir las reglas de la Gramática en nuestra lengua materna, por ejemplo, es probablemente un proceso de aprendizaje inconsciente (*unconscious learning*),[2] en vez de ser similar al aprendizaje de la Gramática de una lengua extranjera en la escuela. La posibilidad de una conducta aprendida de manera no consciente se sostiene, en parte, por la dificultad que tenemos de explicar cómo somos capaces de hacer muchas de las cosas que hemos aprendido a hacer o, incluso, de darnos cuenta de que hemos aprendido una habilidad social. Otro ejemplo de conducta social aprendida de manera no consciente es nuestro conocimiento de la

2. Al abordar la conexión con el aprendizaje se ha alternado el uso de «inconsciente» y «no consciente» para resaltar su distinción respecto de lo completamente inconsciente. *(N. del t.)*

distancia apropiada para mantener una conversación cara a cara con otra persona. Ésta es una norma que casi todo el mundo adquiere en el proceso de socialización, aunque ni la enseñanza ni el proceso de aprendizaje sean conscientes. Caemos en la cuenta de esta regla sólo cuando se nos llama la atención, cuando alguien la incumple o cuando estamos en una Cultura con una regla diferente a la nuestra. Otro punto de apoyo para la tesis según la cual aprender no requiere atención consciente proviene [de la existencia] de formas de vida animal que, de hecho, podemos aprender aun cuando estén desprovistas de conciencia. Puesto que la conducta consciente es una pequeña parte de la actividad humana, podría seguirse que los estudios de la conducta humana no deberían estar limitados a ese terreno.

¿Es, entonces, la conducta voluntaria, en vez de la consciente, la que constituye la acción humana? Normalmente, los filósofos entienden por acciones *voluntarias* aquellas que el agente podría haber dejado de hacer. El concepto de acción voluntaria está conectado estrechamente con la responsabilidad moral. En el contexto moral, *voluntario* significa no coaccionado (*uncoerced*). Podemos ser considerados responsables de acciones que podíamos evitar o cambiar. Aunque esta noción de lo voluntario es adecuada para razonar sobre temas morales, una caracterización diferente de la conducta puede ser más apropiada para la Ciencia Social. En primer lugar, resulta extraño caracterizar la conducta social como voluntaria cuando, como en el caso de mantener la distancia entre hablantes, no somos conscientes de estar comprometidos (*engaged*) en esa conducta. En segundo término, el estudio de la conducta coaccionada está ciertamente dentro del campo de las Ciencias Sociales y ha sido de particular interés para los que intentan entender y explicar actos brutales de crueldad humana.

Normalmente, los interpretacionistas sostienen la importancia de la conducta *dotada de sentido* (*meaningful*) como el objeto relevante para las Ciencias Sociales. Esto es plausible intuitivamente, pero requiere un cuidadoso análisis de qué constituye una conducta dotada de sentido. Creo que entre los interpretacionistas que adoptan este enfoque, Winch es quien delimita con mayor claridad la conducta que desean explicar los científicos sociales (cfr. Winch, 1958). Siguiendo el análisis de Wittgenstein de la conducta lingüística regida por una regla, Winch amplía la noción de regla para incluir las instituciones sociales, las costumbres, las convenciones y así sucesivamente, como también las leyes y las normativas. Del mismo modo que la actividad lingüística se constituye y regula mediante las

reglas de la Gramática, para Winch, así también la conducta social se regula y constituye por reglas sociales. Mientras que autores tan diferentes como Collingwood, Durkheim y Max Weber han subrayado la importancia de las reglas sociales, la perspectiva de Winch acerca de la noción de regla social es lo suficientemente amplia para caracterizar todos los tipos de conducta que los científicos sociales desean explicar: acciones conscientes; acciones voluntarias; acciones que una vez fueron realizadas conscientemente, pero que han llegado a ser automáticas; y acciones que probablemente sean el resultado de un aprendizaje no consciente. Algunas características importantes de la conducta gobernada por reglas, en las que coinciden los naturalistas y los interpretacionistas, son que hay maneras correctas e incorrectas de hacer las cosas de acuerdo con las reglas existentes y que los miembros de una sociedad negocian constantemente —e, incluso, modifican— las reglas de esa sociedad.

Aun cuando sigamos con la tesis de Winch, según la cual es no causal el nexo entre razones y conducta, podemos argumentar que gran parte de la conducta humana gobernada por reglas tiene lugar en ausencia de razones y que las explicaciones causales de esta clase de conducta no incumplen los principios interpretacionistas. Aunque suene paradójico explicar la conducta dotada de sentido sin recurrir a razones, esas explicaciones son habituales y no controvertidas en el campo de la Lingüística histórica. Tras abordar algunos ejemplos en ese terreno, a los que los científicos sociales no les han hecho mucho caso, volveré sobre las sugerencias de Dan Sperber acerca de cómo construir una Ciencia de la Antropología, en la que al menos alguna de las explicaciones sea causal y mecanicista. Mis ejemplos ulteriores están tomados de la Psicología Evolutiva, un campo más problemático. Al terminar, espero haber mostrado que podemos seguir alguna vía para construir un conjunto de explicaciones causales naturalistas en las Ciencias Sociales, sin entrar en la tradicional disputa entre interpretacionistas y naturalistas. Si podemos lograr esto, pienso que podemos comenzar a mostrar cómo la teoría de la causalidad de Wesley Salmon se aplica tanto a las Ciencias Sociales como a las Ciencias Físicas. Puede ser una posición extravagante, si tomamos en serio las observaciones de Paul Humphreys (cfr. Humphreys, 2000), según las cuales «[la teoría de la causalidad de Salmon] nos compromete con el fisicalismo de una manera que podría dar que pensar». Para Humphreys, «esto significa, por ejemplo, que toda la causalidad antropológica y sociológica debe ser aclarada mediante la transferencia de masa-energía, el momento lineal y angular, y así sucesivamente. Y, de este modo, to-

das las explicaciones de esas Ciencias deben darse, en última instancia, en términos de causalidad física». A pesar de su extravagancia, algunos científicos sociales aceptan, sin embargo, esta meta explicativa para su disciplina.

2. La Lingüística histórica

Una tarea preliminar de los lingüistas históricos es explicar el fenómeno del cambio lingüístico. Todas las lenguas vivas están en constante cambio. Aunque sea irregular la inclusión de nuevas características en una lengua, al menos algunos cambios siguen pautas uniformes una vez que han sido introducidos. Esas regularidades son el objeto de investigación de los lingüistas históricos. Son la base de las reconstrucciones parciales de los sistemas de las lenguas prehistóricas de las que no existe información, como la lengua indoeuropea, que es predecesora tanto del castellano como del inglés. Los lingüistas históricos ofrecen explicaciones causales de cómo las lenguas vivas provienen de lenguas prehistóricas que no han sido contrastadas. También intentan explicar las pautas de cambio que se observan en las lenguas vivas. Mucho del trabajo de los lingüistas se concentra en el cambio fonológico, en particular en el cambio en los sonidos vocálicos.

Por ejemplo, algunos cambios de vocales se introducen al parecer en la lengua en cuanto que requieren menor esfuerzo muscular por parte del hablante que la anterior manera de pronunciar la vocal. Los lingüistas denominan este proceso *asimilación*, que significa que un sonido cambia para acomodarse a sonidos cercanos. Los lingüistas creen que los cambios que están causados por asimilación son regulares, lo que comporta que el nuevo sonido, una vez introducido en la lengua,[3] reemplaza al anterior sonido en todos los contextos similares. En algunos casos no ocurre la asimilación esperada, pero los lingüistas dan razón de esas anomalías al mostrar que son el resultado de otros principios regulares del cambio lingüístico, como el préstamo lingüístico o el cambio analógico. Los lingüistas consideran la regularidad de la asimilación como una hipótesis explicativa que es útil también para la reconstrucción lingüística. Aunque la regularidad de la asimilación cuenta con bas-

3. En este contexto se ha traducido *language* por «lengua», en lugar de «lenguaje», por entender que se está hablando del código —la lengua— que plasma una capacidad universal —el lenguaje—. *(N. del t.)*

tante soporte empírico, se la trata como una hipótesis, no como una teoría probada. Es importante para nuestro propósito señalar que el principio de asimilación no se refiere a ninguna intención o decisión por parte del hablante individual para hacer más fácil la pronunciación.

La asimilación es un ejemplo de principio explicativo interno. Acude a características estructurales de la lengua misma para explicar los cambios en la pronunciación. Las explicaciones externas del cambio lingüístico, a diferencia del cambio que está ligado estrechamente a la estructura de la lengua, se basan en lo que ocurre cuando grupos lingüísticos diferentes entran en contacto unos con otros. El castellano que se habla ahora en España contiene muchos elementos que se incorporaron a la lengua durante la ocupación árabe. El castellano que se habla en Iberoamérica muestra la influencia del contacto con las poblaciones indígenas de allí. Los documentos históricos y la investigación arqueológica proporcionan pruebas empíricas (*evidence*) de los contactos entre grupos lingüísticos diferentes. El comercio, la conquista militar y las uniones políticas basadas en matrimonios entre familias gobernantes contribuyen al cambio lingüístico. Esos contactos no introducen sólo gente nueva, bienes, tecnologías, costumbres y sistemas políticos sino también nuevas formas de pronunciar las palabras y palabras nuevas.

Aunque los lingüistas discrepan entre ellos sobre la importancia relativa de las explicaciones internas y externas del cambio lingüístico, los dos tipos de explicación comparten rasgos que son importantes para nuestra visión de la explicación en las Ciencias Sociales:

1) Las explicaciones acuden a causas comunes y procesos causales. Los lingüistas históricos nos dicen que el castellano actual y el italiano moderno se asemejan el uno al otro, porque ambos derivan de una lengua anterior común, el latín vulgar. Los procesos causales que condujeron a su diferenciación incluyen levantamientos sociales y políticos, junto a procesos que complicaron la superación de las barreras geográficas entre Roma y sus provincias. Esos factores causaron alteraciones en la comunicación y permitieron que el latín vulgar, que se hablaba en diferentes lugares del antiguo Imperio romano, se desarrollara de diversas maneras.

2) En su mayor parte, las explicaciones acuden a causas probabilísticas más que a causas deterministas. Por ejemplo, las explicaciones de cómo una lengua proto-indoeuropea, no constrastada, cambió al sánscrito hacen referencia a las frecuencias estadísticas y

a las tendencias de varias vocales de la lengua original que se fusionan en una vocal simple en el sánscrito. Muchas explicaciones internas del cambio lingüístico se refieren a las fuerzas condicionantes, tales como la asimilación —que propicia el cambio de un sonido particular— junto a otros factores de contrapeso, que tienden a prevenir el cambio o a hacer que sea incompleto. Los levantamientos políticos y otras causas externas del cambio lingüístico son causas probabilísticas.

3) Raramente las explicaciones en la Lingüística histórica apelan a razones individuales, incluso cuando las decisiones humanas documentadas son relevantes para la situación lingüística. Los lingüistas históricos son conscientes de que las decisiones individuales pueden tener consecuencias importantes para el cambio lingüístico. Las decisiones militares y políticas de los líderes pueden determinar el éxito o el fracaso de las conquistas o de los acuerdos comerciales y, así, *a fortiori*, pueden causar o impedir ciertos tipos de cambios lingüísticos. Los deseos de individuos particulares de modificar su *status* imitando al rico, al famoso o al poderoso puede causar cambios en el vocabulario y en cómo se pronuncian las palabras. Las decisiones de los individuos de dedicarse al comercio origina cambios en el lenguaje, ya que los hombres de negocios buscan una base lingüística común para sus negociaciones. Sin embargo, las creencias y los deseos de los individuos apenas son visibles en las explicaciones del cambio lingüístico de los lingüistas históricos. Se pueden dar varias razones que aclaran ese abandono o supresión de las creencias y deseos individuales. Obviamente, obtener información sobre las decisiones individuales es casi imposible para muchos de los periodos prehistóricos estudiados por los lingüistas. Además, las investigaciones lingüísticas se enfocan, no obstante, hacia los procesos de cambio interno y externo, en lugar de hacia los mudables motivos de los agentes de esos procesos. Cómo se introduce una innovación en el lenguaje parece menos importante que los procesos mediante los cuales se mantiene, si es que lo hace. Casi se podría decir que, al igual que los biólogos evolutivos se centraron en los procesos de selección en vez de en las mutaciones —que son la última fuente de cambio—, así también los lingüistas se han ocupado de lo que ocurre tras haber incorporado un nuevo ítem a la lengua. Las decisiones individuales son tratadas como accidentes históricos. Los así llamados «accidentes» no son, desde luego, incausados. Pero sus causas no son realmente relevantes para el trabajo de los lingüistas históricos, que estudian los procesos del cambio lingüístico. Esos accidentes históricos son las

fuentes de la innovación. Sin embargo, la nueva característica, una vez introducida en la lengua, puede ser adoptada por razones internas o externas, o puede ser que no cale. A veces, las nuevas características se adentran en un área, pero se evita su propagación a otros hablantes de esa lengua, por el aislamiento político o geográfico. Cuando sucede esto, las innovaciones tienden a acumularse y a experimentar nuevos cambios, así que, al final, los subgrupos de una población antiguamente unida llegan a estar separados lingüísticamente en la misma medida en que se diversifican los subgrupos de las especies biológicas, hasta que no pueden interrelacionarse por más tiempo.

Los lingüistas históricos también reconocen que, mientras que una nueva incorporación a la lengua puede ser el resultado de una decisión consciente de algún individuo, el cambio lingüístico actúa en su mayor parte en ausencia de una conciencia expresa. Los procesos que mencionan los lingüistas para dar razón del cambio, tales como la asimilación, no son normalmente procesos de los que sean conscientes los hablantes. El descubrimiento de la regularidad del cambio de sonido fue un importante logro empírico, que permitió a los lingüistas explotar las correspondencias entre el sonido y el significado (*meaning*), pero esa regularidad y las explicaciones que le subyacen —como la mayoría de las explicaciones en la Lingüística histórica— no hacen ninguna referencia a la intención, motivos o razones de los hablantes. Todo esto es cierto, aun cuando la conducta lingüística esté dotada de sentido (*meaningful*) y esté gobernada por reglas (vid. M. H. Salmon, 1996).

3. **Antropología**

¿Es posible estudiar la conducta social (otra que el cambio lingüístico) usando técnicas similares a aquéllas de los lingüistas históricos? En *Sobre el conocimiento antropológico* (1985) y *Para explicar la Cultura* (1996), Dan Sperber ofrece un enfoque de la Antropología que se parece a la Lingüística histórica en que fija su atención en la transmisión de componentes culturales y en los procesos de cambio cultural, en vez de [centrarse en] la interpretación de las creencias y los deseos de los individuos. Reconoce la importancia del trabajo interpretativo en la Antropología, pero piensa que es imposible formular generalizaciones científicas útiles sobre la base de las interpretaciones. A este respecto, está de acuerdo con la crítica interpretacionista habitual acerca de la posibilidad de la Ciencia So-

cial. Sin embargo, cree también que podemos separar los aspectos interpretativos de la Antropología de aquellas características de la disciplina que se prestan al estudio científico. Según Sperber, el objeto de la Antropología científica es la distribución de las creencias o cómo algunas creencias llegan a formar parte de la Cultura.

Algunas creencias —o representaciones mentales— son privadas, cosas de índole efímera. Puede que nunca sean transmitidas a otra persona o incluso, si lo son, puede que no lleguen nunca a estar ampliamente difundidas o lo suficientemente bien establecidas como para ser consideradas parte de una Cultura. Otras creencias, expresadas como descripciones orales o escritas, son ampliamente compartidas, cosas de carácter público. Estas últimas creencias —o representaciones públicas— son culturales. Sperber cree que una Ciencia de la Antropología debería ofrecer explicaciones causales naturalistas (fisicalistas) de cómo las representaciones —o, de modo más preciso: cómo diversos tipos de representaciones— llegan a difundirse, es decir, cómo llegan a convertirse en Cultura. Llama a este estudio «la epidemiología de las representaciones». Para forzar la analogía con la Lingüística histórica, podríamos decir que Sperber cree posible construir una genuina Ciencia de la Antropología que explicaría causalmente el desarrollo y el cambio de las Culturas y lo haría así, en su mayor parte, de una manera similar a las explicaciones internas del cambio lingüístico de los lingüistas. El problema del contenido de las creencias volvería a ser un estudio válido dentro de una rama interpretacionista de la Antropología; la disciplina a la que Sperber menciona como la Etnología. La rama científica de la Antropología —la Antropología misma— se ocupa de la difusión de las ideas, en lugar de su contenido. Al igual que los lingüistas, que estudian factores internos, como el cambio de un sonido para ajustarse a sonidos cercanos, y factores externos, como las conquistas militares, que transforman las lenguas de manera más o menos regular, del mismo modo Sperber cree que los antropólogos deberían estudiar los factores psicológicos y ecológicos que contribuyen a la difusión de las representaciones. Ejemplos de factores psicológicos son el atractivo de las representaciones y la facilidad con la que pueden ser memorizadas. Los ejemplos de factores ecológicos incluyen la disponibilidad de un mecanismo externo para almacenar la memoria, tales como la escritura, y el entorno institucional de una sociedad (Sperber, 1996, pp. 85 y 96). Como en los casos de los factores internos y externos del cambio lingüístico, los factores psicológicos y ecológicos pueden contrarrestarse o reforzarse el uno al otro; a menudo, los dos tipos de factores están implicados en una explicación del cambio cultural.

Sperber describe dos fuentes principales de las creencias: las percepciones —y las inferencias que se basan en ellas— y la comunicación. Todas nuestras creencias dependen de la comunicación, en cuanto a la adquisición de los conceptos necesarios para caracterizar las percepciones. Pero algunas creencias, aquellas que Sperber denomina «intuitivas», dependen para su distribución de una amplia gama de experiencias perceptivas y de muchas experiencias comunicativas convergentes. Esas creencias se adquieren inconscientemente en su mayor parte. Forman la base de nuestra visión del mundo de sentido común y de nuestra habilidad para comunicarnos los unos con los otros. Según Sperber, esas representaciones ampliamente compartidas no sólo determinan gran parte de nuestra conducta sino que también son la base de las «creencias reflejas» (*reflective beliefs*). Estas últimas creencias no se fundan en la percepción sino que dependen casi por entero de la comunicación para su distribución (Sperber, 1996, p. 94). Las creencias reflejas incluyen tanto los mitos como las teorías científicas de alto nivel. Pueden ser comprendidas por completo, como la comprensión por los matemáticos del Teorema de Incompletud de Gödel, o pueden ser mal entendidas, pero sin embargo creídas, como cuando son aceptadas a tenor de la confianza en una autoridad. Para ilustrar cómo diferentes factores afectan la distribución de las creencias reflejas, Sperber trata tres ejemplos: la diseminación de los mitos en una sociedad iletrada; la difusión de las creencias políticas, tales como la creencia de que todos los seres humanos han sido creados iguales, y la distribución de una compleja creencia matemática, a saber: la prueba del Teorema de Incompletud de Gödel. La transmisión de un mito, tal como «caperucita roja», depende principalmente de su atractivo y su facilidad para ser memorizado, los dos son factores psicológicos (cognitivos). Mitos como ése se encuentran en cada Cultura, y no parece que su supervivencia dependa de la existencia de instituciones especializadas. Por el contrario, la transmisión de una creencia política depende, principalmente, de su relevancia para las instituciones políticas existentes, que son factores ecológicos. Sin un nexo con alguna institución que asuma, al menos, una forma limitada de gobierno democrático, la creencia según la cual todos los seres humanos han sido creados iguales no hubiera conseguido llegar a convertirse en una creencia cultural. La transmisión de la creencia en la prueba de Gödel difiere de los mitos y de las creencias políticas, porque depende en gran medida a la vez de factores cognitivos y de la existencia de instituciones que comuniquen esas formas abstractas de conocimiento. Para aquellos que

entienden del todo la prueba, los factores cognitivos fuerzan su aceptación, pero para aquellos que no la entienden plenamente depende de factores institucionales, tales como la aceptación de esa creencia por los expertos.

La mayoría de las explicaciones internas del cambio fonético que dan los lingüistas históricos son explicaciones causales naturalistas. Esas explicaciones apelan a características estructurales del lenguaje más que a instituciones sociales o políticas. Los sonidos del habla, en cuanto tales, son cosas físicas, se producen por procesos causales físicos mecánicos bien conocidos (colocar la lengua y los labios de una manera determinada, emitir un soplo de aire, y así sucesivamente). Se puede medir la cantidad de esfuerzo físico requerido para producir sonidos vocálicos en diversos contextos. Un principio de conservación del esfuerzo no requiere apelar a la intención o el deseo, y parece susceptible de quedar enmarcado en términos fisicalistas (*physicalistic*). Con el conocimiento de las causas físicas de esta índole, la explicación de un cambio vocálico en términos de asimilación es, aparentemente, una explicación causal naturalista. Algunas explicaciones del cambio lingüístico incluyen también causas físicas. Por ejemplo, el aislamiento geográfico de algunos miembros de un grupo lingüístico es una barrera física para la comunicación libre. Cuando los subgrupos de una comunidad lingüística llegan a estar separados físicamente uno respecto de otro, la comunicación puede disminuir o cesar. Entonces, la lengua que, una vez fue común, se desarrolla de forma diferente en los grupos separados. Otros cambios lingüísticos externos, como los esfuerzos de individuos de clase baja por imitar la pronunciación de un grupo de élite, aun cuando es posible que no dependan de las creencias conscientes y de los deseos de los individuos, no por ello resultan adaptables de modo obvio a una explicación naturalista (fisicalista). Por esta razón, en parte, algunos lingüistas prefieren las explicaciones internas del cambio lingüístico a las externas. Está claro, sin embargo, que al menos algunas explicaciones de la Lingüística histórica se acercan a los estándares señalados por Salmon y descritos por Humphreys.

En el caso de la Antropología, Sperber hace una serie de propuestas sobre cómo se pueden construir las explicaciones causales naturalistas. En primer lugar, puesto que la transmisión de las representaciones es el tema que tenemos entre manos, ha de mostrar que las representaciones mismas son objetos físicos. Las representaciones culturales no son un gran problema, ya que son oraciones habladas o escritas, inscripciones, comportamientos, pinturas, edi-

ficios, y así sucesivamente. Todas ellas son cosas físicas, del mismo modo que los sonidos son objetos físicos ¿Qué ocurre con las representaciones mentales particulares? La concepción de Sperber es que son configuraciones del cerebro, y también objetos físicos. Desde luego que esta concepción no es compartida, en general, por los no materialistas. Los neurocientíficos, sin embargo, están comprometidos con ella. Aunque la naturaleza exacta de las relaciones entre los diversos estados cerebrales y la representación mental es un tema de investigación intensa y problemática, parece que hace falta una teoría que conecte los estados cerebrales y los estados mentales, para que sea plausible una exposición causal fisicalista de las Ciencias Sociales.

Si podemos aceptar la naturaleza física de las representaciones mentales, el siguiente paso para proporcionar explicaciones naturalistas de la transmisión de ideas es ofrecer una exposición fisicalista de los mecanismos que transforman las representaciones mentales privadas en representaciones culturales. Algunos autores —el más famoso es R. Dawkins, que acuñó el término «meme» para describir la unidad de Cultura (Dawkins, 1982)— han propuesto la selección natural como mecanismo. Sperber considera esta idea pero la rechaza, sobre todo porque las representaciones no se reproducen a sí mismas del modo como lo hacen los genes. Mientras que los genes se reproducen a sí mismos, las ideas se transmiten, y la transmisión casi siempre comporta algún grado de modificación. Para ver que esto es así, considérese lo que ocurre con un mensaje simple cuando se pasa de una persona de una oficina a otra, de la segunda persona a la tercera, y así sucesivamente. Si la cadena de transmisión no incluye más que unas pocas personas, es casi seguro que el mensaje se cambia por el camino, a veces hasta el extremo de apenas parecerse a la versión original. Así, aunque no siempre los genes se reproduzcan con éxito —se dan mutaciones—, la norma es que haya copias idénticas de genes, y esta clase de estabilidad es necesaria para que la selección funcione. Por el contrario, cuando se transmite una idea, la norma es la alteración. La replicación de representaciones es posible como caso límite —por ejemplo, la reproducción de copias mecánicas o electrónicas de un libro impreso—, pero esto son excepciones a la regla.

Sperber rechaza también el «modelo de influencia» para la transmisión de elementos culturales. Ese modelo presenta las nuevas ideas culturales como resultado de entradas de diferente origen, que pueden variar en cuanto a cómo sea de intensa la influencia que ejercen sobre el elemento transmitido. A menudo, se recurre a este

modelo en las Ciencias Sociales, al igual que en las explicaciones corrientes acerca de la forma de las representaciones públicas, como los libros o las obras de arte. Consideremos, por ejemplo, que una estudiante escribe un trabajo de investigación sobre el desarrollo de la agricultura en Iberoamérica. Supongamos que la estudiante trabaja con un preeminente profesor marxista, pero tiene también un conjunto de evaluadores no marxistas en su tribunal. Una evaluación de la Tesis podría versar sobre los variables niveles de intensidad de las influencias que confluyen en el producto final. Una gran parte de la Historia de las ideas se ocupa, por supuesto, de trazar las influencias como un modo de comprender y explicar las formas culturales. El argumento principal de Sperber contra este enfoque es que, como en la replicación, no tiene en cuenta la genuina novedad en los elementos transmitidos. Una vez que se ha dado un peso a las diversas influencias, se las trata como entradas que terminan en un resultado que está completamente determinado. Cree, por el contrario, que las representaciones mentales de los elementos transmitidos pueden contener más información que la representada mediante la síntesis de los aspectos de entrada (Sperber, 1996, p. 106).

La concepción personal de Sperber es que las representaciones *se transforman* a sí mismas, como fruto de un proceso de construcción cognitiva. Hace falta, desde luego, algún grado de semejanza (*resemblance*) entre las representaciones mentales individuales, si han de ser consideradas como representaciones de la misma Cultura; pero el proceso por el que se transmiten las representaciones no es ni la producción de una copia de sí mismo ni tampoco la producción de un compuesto obtenido a partir de diversas representaciones influyentes.

«Atracción» se llama lo que Sperber elige como mecanismo. La atracción comporta la convergencia de los propios «procesos cognitivos y afectivos y de aquellos otros de muchas otras personas que tienden hacia algún tipo atractivo de perspectivas dentro del amplio campo de las posibles concepciones» (*ibidem*). La atracción es un concepto estadístico, y no conlleva ningún objeto físico que atraiga a las representaciones similares. Por ejemplo, hay versiones mejores y peores de un mito, como «caperucita roja», pero eso no significa que haya copias más o menos fidedignas de alguna primera versión. En cambio, las narraciones del cuento tienden a converger en una versión que es psicológicamente atractiva, en el sentido de ser más fácil de recordar, al tener muchos menos «cabos sueltos», y otros rasgos de esta índole. Los factores implicados en la atracción

son tanto psicológicos como ecológicos, es decir, que el entorno proporciona las entradas a los individuos y su organización mental determina qué entradas se procesan, cómo se procesan y cómo afectan a la conducta. En el enfoque de Sperber, los cambios culturales a gran escala son, en última instancia, el resultado de esos cambios a pequeña escala, como podría ocurrir con las variaciones en cuanto a la convergencia de la mejor versión de «caperucita roja» o la manera correcta de pronunciar una vocal en un contexto dado.

4. La Psicología evolutiva

Aunque Sperber rechaza la evolución como mecanismo para el cambio cultural, sus puntos de vista tienen bastante en común con el trabajo de psicólogos evolutivos como J. Tooby y L. Cosmides (Tooby y Cosmides, 1992). Los psicólogos evolutivos están comprometidos con el punto de vista según el cual, del mismo modo que el cuerpo humano se ha configurado mediante la presión evolutiva, así lo ha hecho también la mente humana. Esto no significa que cada rasgo humano se haya formado por selección natural, sino que los diversos componentes mentales (módulos) se han configurado para tareas especiales aunque definidas en un sentido amplio, tales como usar un lenguaje hablado. Los estudios del cerebro que usan técnicas de representación (*imaging techniques*) ofrecen un soporte a los módulos mentales (*mental modules*), porque han mostrado conexiones regulares entre ciertas clases de estimulación mental y la actividad de áreas específicas del cerebro. La Psicología evolutiva tiene la tarea de demostrar la ventaja selectiva de las diversas características humanas y, también, de dar razón de los rasgos disfuncionales o claramente desventajosos. La ventaja selectiva de tener lenguaje oral es bastante obvia, mientras que la ventaja selectiva de otras actividades humanas admirables pero costosas, tales como tocar muy bien un instrumento musical o dedicarse a la Matemática de alto nivel, no está tan clara. Del mismo modo que los biólogos evolucionistas tienen la onerosa tarea de explicar cómo los alelos nocivos, como la malaria, persisten en una población, así los psicólogos evolutivos tienen la ardua tarea de explicar la persistencia de la conducta en apariencia disfuncional, como el infanticidio, el suicidio, y cosas semejantes.

Los psicólogos evolutivos desean explicar también cómo los humanos pueden ser distintos culturalmente, a pesar de ser muy similares en su conformación psicológica. Lo hacen acudiendo a señalar

la importancia de los factores ambientales. El explicar el cambio cultural en términos evolutivos conlleva dificultades especiales, que no están presentes al explicar el cambio biológico. El fracaso de las unidades culturales para replicarse ya ha sido tratado. Además, la magnitud de tiempo del cambio evolutivo es muy grande, comparada con la magnitud para el cambio cultural. Este problema, en parte, se responde al argumentar que muchos módulos formados evolutivamente están diseñados para proporcionar respuestas «facultativas» (*facultative*) a las presiones ambientales. Del mismo modo que la piel se broncea en diversos grados, cuando se expone al sol, las respuestas culturales a los estímulos ambientales pueden ser muy rápidas. Hay, probablemente, en el cerebro de todo humano un módulo de procesamiento del lenguaje, pero el (primer) lenguaje que adquiere el ser humano depende del lenguaje disponible en su entorno.

Aunque son enormes las dificultades que afronta una Ciencia de la Psicología evolutiva plenamente desarrollada, creo que se han hecho algunos progresos al dar con los retos que se acaban de mencionar. El programa está comprometido con la formulación y contrastación de hipótesis empíricas, dentro del amplio marco teórico de la Teoría evolutiva, y para ofrecer explicaciones que acudan a mecanismos físicos que den razón de la conducta humana. Sobre éste último aspecto, el programa está comprometido con un tipo de explicación causal coherente con el análisis de Salmon. Veamos, brevemente, dos ejemplos que contrastan las hipótesis usadas para explicar la conducta humana en términos de la promoción de la supervivencia de los propios genes a través de favorecer a las personas relacionadas genéticamente con ellos. Ninguno de los ejemplos apela a las intenciones de los agentes que intervienen. El primer ejemplo es la explicación de por qué en algunas sociedades, el hombre se responsabiliza en menor grado del bienestar de los hijos de su propia esposa, pero presta especial atención a los hijos de su hermana. Esta pauta de comportamiento se denomina «herencia matrilineal» (*matrilineal inheritance*). La hipótesis evolutiva invocada para dar razón de esto es que los varones gastan sus recursos en los hijos en proporción a su (probable) grado de relación genética con ellos. Se podría suponer entonces que las sociedades en las que los varones no asumen primariamente la responsabilidad sobre los hijos de sus esposas son sociedades en las que las mujeres son proclives a entablar relaciones sexuales extramaritales. En esas sociedades, la paternidad del marido está en duda. Si el varón presta considerable atención a los hijos de su esposa, puede estar apoyando a un niño que no está relacionado genéticamente con él. Sin em-

bargo, puesto que la maternidad no se pone casi nunca en duda, puede estar seguro de que él y su hermana comparten los mismos genes. Los hijos de su hermana llevan también algunos de esos genes. Cuidando los hijos de su hermana, puede estar seguro de invertir en niños que portan alguno de sus genes. Varios estudios empíricos, que incluyen uno importante de J. Hartung (Hartung, 1985), han sido diseñados para contrastar la hipótesis de que la herencia matrilineal y la probabilidad de la paternidad están correlacionadas de la manera descrita. Los resultados de todos ellos sustentan la hipótesis (Gaulin y McBurney, 2001, pp. 346-349).

El segundo ejemplo conlleva la investigación de homicidio desde el punto de vista de la Psicología Evolutiva (Daly y Wilson, 1988). Centrándose en los datos del censo de 1970 de la ciudad de Detroit, investigaron homicidios domésticos. En situaciones del hogar que afectan tanto parientes (consanguíneos) como no parientes (esposos e hijastros), encontraron que, a cada tipo de no-pariente, se le mata más a menudo de lo que se esperaba por casualidad (*chance*), y que en cada tipo de pariente era menos frecuente de lo esperado por casualidad. En promedio, según el sumario de este trabajo en Gaulin y McBurney, «los corresidentes no parientes tienen once veces más riesgo de homicidio doméstico que los corresidentes parientes» (Gaulin y McBurney, 2001, p. 324).

Ambos ejemplos permiten suscribir la concepción según la cual estamos diseñados por la evolución para ser más proclives a ayudar y a tender a dañar menos a aquellos que están relacionados genéticamente con nosotros, en comparación con aquellos que no lo están. Esos resultados son ofensivos para alguna gente, porque creen que sugieren que esa conducta es inevitable, o incluso que «nadie es culpable». Les preocupa que, al mostrar que una conducta es el resultado de la presión evolutiva, se arrojen dudas sobre la actuación moral humana. Esas interpretaciones, sin embargo, son incorrectas. La conducta está conformada a la vez por la psicología de los individuos y por el ambiente, ninguno de esos factores es inmune al cambio. Los rasgos evolutivos no están fijados, particularmente aquellos que son respuestas facultativas al entorno. Cambiando el entorno se puede modificar la conducta. El estudio sobre el homicidio, por ejemplo, muestra que los no parientes son más vulnerables que los parientes en las situaciones del hogar. A la luz de ese conocimiento, los agentes morales pueden modificar los entornos domésticos para hacer menos vulnerables a los no parientes. La modificación psicológica también es una opción, y ha mostrado tener éxito en algunos programas de consulta para aquellos casos propensos a los abusos domésticos.

5. Conclusión

En este trabajo he procurado arrojar luz sobre algunas propuestas sobre la explicación científica de la conducta humana. Que esas propuestas sean coherentes con el enfoque de Salmon de la causalidad en el mundo físico es un punto a su favor. Apreciar este intento de introducir a las Ciencias Sociales en la línea de las otras Ciencias no desmerece en ningún sentido la importancia del quehacer interpretativo, como el llevado a cabo en la Etnografía tradicional. Las intuiciones sobre la conducta humana obtenidas por los estudios interpretativos —así como por el trabajo literario— son significativas. Sin embargo, la posibilidad de explicaciones causales de la conducta humana tiene una enorme significación práctica, como también filosófica, porque el conocimiento de las causas es crucial para modificar la conducta que sea peligrosa y destructiva.

La necesidad de ese conocimiento causal y de ese control es patente ahora incluso para los humanistas. Véase, por ejemplo, el reciente trabajo de Jonathan Glover: *La Humanidad: Una historia moral del siglo veinte* (1999). Glover realiza un recuento y analiza con escalofriante detalle los desastres morales del siglo pasado, incluyendo el genocidio en la Alemania nazi y Ruanda, el asesinato en masa de civiles en Hiroshima y My Lai, junto con el terrorismo y los asesinatos en masa en la Rusia estalinista y en Camboya. Glover, cuyas credenciales como humanista son impecables, y cuyo lema de comienzo está tomado de la autobiografía de Collingwood, comienza y termina su libro con un alegato en favor de la comprensión de las causas que nos permitirán controlar esa mala conducta. Al final de su introducción, pide disculpas por el oscuro tema del libro, pero pasa luego a decir que «necesitamos mirar dura y claramente a algunos monstruos dentro de nosotros. Pero esto es una parte del proyecto de enjaularlos y domesticarlos» (Glover, 1999, p. 7). Sus comentarios finales son: Para evitar nuevos desastres, necesitamos controles políticos a escala mundial. Pero la política no es toda la historia. Hemos experimentado los resultados de la Tecnología al servicio de la cara destructiva de la psicología humana. Algo hay que hacer respecto de esta combinación fatal. Los medios para poner en práctica la crueldad y para llevar a cabo asesinatos en masa [ya] han sido desarrollados por completo. Es demasiado tarde para detener la Tecnología. Es a la Psicología a lo que deberíamos volver ahora (cfr. Glover, 1999, p. 414).

Referencias bibliográficas

COLLINGWOOD, R. G. (1946, 1956), *The Idea of History*, Oxford University Press, Oxford.

DALY, M. y WILSON, M. (1988), *Homicide*, Aldine de Gruyter, N. York.

DAWKINS, R. (1982), *The Extended Phenotype*, Oxford University Press, Oxford.

DURKHEIM, E. (1897, 1951), *Suicide*, The Free Press, N. York.

GAULIN, S. y MCBURNEY, D. (2001), *Psychology: An Evolutionary Approach*, Prentice Hall, Upper Saddle River, N. Jersey.

GLOVER, J. (1999), *Humanity: A Moral History of the Twentieth Century*, Yale University Press, N. Haven y Londres.

HARTUNG, J. (1985), «Matrilineal Inheritance: New Theory and Analysis», *Behavioral and Brain Sciences*, v. 8, pp. 661-670.

HUMPHREYS, P. (2000), «Review of *Causality and Explanation* by W. C. Salmon», *Journal of Philosophy*, v. 97, n. 9, pp. 523-527.

MCINTYRE, L. (1999), «Davidson and Social Scientific Laws», *Synthese*, v. 120, n. 3, pp. 375-394.

MILL, J. S. (1874), *A System of Logic* (8.ª ed.), M. Harper and Bros., N. York.

SALMON, M. H. (1996), «Causal Explanations of Linguistic Behavior», *Theoretical Linguistics*, v. 22, pp. 195-205.

SPERBER, D. (1985), *On Anthropological Knowledge*, Cambridge University Press, Cambridge.

SPERBER, D. (1996), *Explaining Culture*, Blackwell, Oxford.

TOOBY, J. y COSMIDES, L. (1992), «The Psychological Foundations of Culture», en BARKOW, J., COSMIDES, L. y TOOBY, J. (eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford University Press, N. York, pp. 19-136.

WINCH, P. (1958), *The Idea of a Social Science*, Routledge and Kegan Paul, Londres.

PARTE V

EXPLICACIONES TELEOLÓGICO-CAUSALES Y EXPLICACIONES PROGRAMÁTICAS

Capítulo 8

FINALIDAD Y CAUSALIDAD EN LAS EXPLICACIONES CIENTÍFICO-SOCIALES. ANÁLISIS DEL ENFOQUE DE R. TUOMELA

Rafaela García Elskamp

1. Contraposiciones metodológicas

Uno de los problemas de fondo de las Ciencias Humanas y Sociales es la articulación entre finalidad y causalidad. Es una cuestión que engarza con otra bien conocida: ¿forman los Saberes Humanos y Sociales un todo con otras disciplinas o, por el contrario, difieren netamente entre sí, llegando a ser saberes esencialmente distintos? Esta polaridad entre unidad y diversidad de la Ciencia, que late tras el binomio finalidad-causalidad, puede ser vista a tenor de la dualidad *Erklären-Verstehen*. «Explicación» y «comprensión» son términos que expresan una dualidad metodológica de gran importancia. Su uso ha variado de unos autores a otros,[1] pero siempre ha servido para transmitir la idea de una tensión metodológica.

Habitualmente, *Erklären* aparece vinculado al estudio de causas y asociado a la posibilidad de leyes o regularidades estables. Se relaciona, en principio, con algo que puede ser repetido o reproducido y, por tanto, es predecible. Explicar el porqué de un fenómeno puede propiciar, en efecto, la posterior predicción de un suceso futuro. Un suceso que, en principio, acaecerá de nuevo si vuelve a repetirse la causa. En cambio, *Verstehen* no queda dentro del ámbito gobernado por leyes o regularidades. Comprender implica algo más bien cualitativo que cuantitativo: intenta abarcar al hombre, sus re-

1. Recientemente se han distinguido siete enfoques diferentes de esta dualidad; cfr. González, W. J., «From *Erklären-Verstehen* to *Prediction-Understanding*: the Methodological Framework in Economics», en Sintonen, M., Ylikoski, P. y Miller, K. (eds.), *Realism in Action*, Kluwer, Dordrecht, 2002, pp. 38-48; en especial, pp. 34-37.

laciones interindividuales y sus vínculos con la realidad en general.[2] Afronta así eventos que atañen al dominio de la libertad humana.

Este binomio metodológico encuentra su origen en W. Dilthey y en su interés por resaltar la estructura específica del objeto de lo que denominó *Geisteswissenschaften.*[3] Para Dilthey, la comprensión era un modo de captar intuitivamente los elementos subjetivos presentes en las acciones humanas. Básicamente desde esta perspectiva se intenta conocer el sentido y la finalidad de los hechos. *Verstehen* lleva a entender que, al hablar sobre el hombre y la sociedad, la mirada no ha de estar dirigida hacia leyes más o menos invariantes, sino que primeramente intenta hacer inteligible la actividad humana indagando sobre su sentido y finalidad. El término «comprensión» ha ido abarcando diversos aspectos según la caracterización particular de cada autor: G. E. M. Anscombe, Ch. Taylor, G. H. von Wright,... En general, como ya se ha señalado, se enfrenta a la explicación causal: como origen de los eventos o hechos busca algo que no sean causas (en la acepción que éstas tienen en las Ciencias de la Naturaleza).

La *comprensión* es una aproximación a la realidad humana, entendida ésta como un elemento surcado por la libertad y la variabilidad de cada agente individual o social. Al buscar la racionalidad de los hechos humanos se entra en el mundo de las intenciones, motivos, razones, propósitos, interpretaciones o valores desarrollados por los grupos sociales. Se sitúa ante el amplio campo de la experiencia humana. Sin embargo, en el campo de las Ciencias Humanas y Sociales también se ha intentado encontrar una posibilidad de aplicación práctica, que permita dominar y, en lo posible, predecir el decurso de las situaciones humanas.

Frente a esto, los partidarios más estrictos del *Erklären* dejan de lado los elementos cualitativos y teleológicos y se centran en los causales y cuantitativos. Su uso aparece vinculado al estudio de causas, normalmente causas naturales. De hecho, la explicación tiene por campo característico la Naturaleza, es decir, algo que puede ser experimentado y es susceptible de expresión matemática. Se interesa por fenómenos que, normalmente, son reproducibles, lo cual facilita que sean predecibles.

2. Cfr. GONZÁLEZ, W. J., «Caracterización del objeto de la Ciencia de la Historia y bases de su configuración metodológica», en GONZÁLEZ, W. J. (ed.), *Acción e Historia. El objeto de la Historia y la Teoría de la Acción*, Publicaciones Universidad de A Coruña, A Coruña, 1996, pp. 75-76.

3. DILTHEY, W., *Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studiem der Gessellschaft und der Geschichte*, en DILTHEY, W., *Gesammelte Schriften*, Teubner y Vandenhoeck-Ruprecht, Stuttgart y Gotinga, 5.ª edic., 1962.

Esta caracterización general de la dualidad mencionada queda remarcada por von Wright, especialmente cuando escribe que «la comprensión se encuentra vinculada con la intencionalidad de una manera en que no lo está la explicación. Se comprenden los objetivos y propósitos de un agente, el significado de un signo o de un símbolo, el sentido de una institución social, o de un rito religioso».[4]

Según un modo de presentación habitual, en la explicación causal se recurre a un origen que hace inteligibles los fenómenos en cuanto efectos de aquel origen. Esta manera de explicar plantea una relación necesaria, o por lo menos «nómica», entre la causa y su efecto. En cambio, la explicación teleológica recurre a la finalidad: analiza la acción en función del fin que se pretende. Este fin queda siempre recogido en la intención de la acción. Así, esta última aproximación a la realidad pretende comprender la acción, y esto sólo es posible atendiendo tanto a los motivos como a las razones que el agente da de su propia acción.[5]

Dejando de lado esta conocida controversia metodológica, nos preguntamos ahora cómo se genera de hecho una acción, ya que las acciones de los individuos no son acciones aisladas. Las acciones se encuentran interconectadas, al igual que lo están los sujetos que las producen. Las personas interactúan, de modo que las acciones que unos llevan a cabo propician las de aquellos otros a los que alcanza la acción inicial. Aquí habría múltiples aspectos a considerar: tiempo, habilidad, tipo de generación... Todos ellos son rasgos públicos u observables; son características susceptibles de ser analizadas desde el exterior del agente, con independencia de que estemos ante acciones sociales o individuales. Obviamente, junto a éstos se encuentran los actos mentales.

Las coordenadas del problema del patrón explicativo de las Ciencias Humanas y Sociales implican —a mi juicio— atender a los rasgos constitutivos de la acción humana. Ésta es, en efecto, el objeto propio de estas disciplinas, que pueden ser vistas como «Ciencias de la Acción», tal como ha hecho, entre otros, von Wright.[6] En las Ciencias Humanas y Sociales encontramos las acciones libres de

4. WRIGHT, G. H. VON, *Explanation and Understanding*, Cornell University Press, Ithaca, 1976, p. 6.

5. Sobre la distinción entre «motivos» y «razones», cfr. RESCHER, N., *Razón y valores en la Era científico-tecnológica*, Paidós, Barcelona, 1999.

6. Cfr. WRIGHT, G. H. VON, *Explanation and Understanding*, *passim*, y WRIGHT. G. H. VON, «Action and Explanation of Behaviour», *Psykologia*, v. 1, (1966), pp. 132-135. Véase también TUOMELA, R., *Human Action and Its Explanation*, Reidel, Dordrecht, 1977; y MANNINEN, J. y TUOMELA, R. (eds.), *Essays on Explanation and Understanding*, Reidel, Dordrecht, 1985.

los hombres como realidad básica, mientras que en las Ciencias de la Naturaleza los fenómenos empíricos aparecen desprovistos de esta característica. De ahí que los factores de unidad y diversidad de la Ciencia puedan ser analizados a tenor de la Teoría de la Acción. Por eso se apuntan aquí elementos específicos de los saberes humanos y sociales atendiendo a la Teoría de la Acción.[7]

2. Las acciones como base de las Ciencias Humanas y Sociales

Un sistema explicativo de la acción que pretenda ser completo debe abarcar tanto los antecedentes de la misma —los actos mentales— como el movimiento corporal —o su omisión—, aquello que es observable por un agente exterior a la acción. En esta línea, L. Wittgenstein se planteaba: «¿por qué quiero yo contarle a él tanto mi intención como lo que hice? No porque la intención también fuera algo que entonces ocurría, sino porque quiero comunicarle a él algo acerca de mí mismo, que va más allá de lo acontecido en aquel momento».[8] Las acciones no se reducen a algo observable, este aspecto es, simplemente, la punta del iceberg. La parte material de la acción viene sustentada por su *intención* (antecedente o elemento no material) y, por lo tanto, no observable desde el exterior. Estos antecedentes dan el sentido a la acción, además ofrecen una valiosa información sobre el estado de ánimo en el que se encuentra el agente. A esto hay que añadir la explicación de la relación que se da entre ambos planos.

En el caso de las acciones sociales, entendiendo por tales aquellas realizadas por más de un agente, el modelo explicativo de la acción intencional debe incluir también las *creencias mutuas* que los agentes actuantes poseen acerca de las actitudes comunes que mueven a todos los participantes en la acción. Es decir, han de incluir las que conforman a un grupo de modo colectivo, así como la creencia individual de que los demás comparten las mismas opiniones. Se trata en definitiva, de las creencias comunes compartidas que proporcionan fuerza y unidad al grupo.

7. A este respecto, entre los estudios en nuestro país destaca el libro, ya citado, sobre *Acción e Historia. El objeto de la Historia y la Teoría de la Acción*. Desde el punto de vista de la propia estructura de la Teoría de la Acción hay que resaltar el conjunto de trabajos del volumen GONZÁLEZ, W. J. (ed.), *Action Theory—Teoría de la Acción*, número monográfico de *Daimon*, v. 3, (1991).

8. WITTGENSTEIN, L., *Philosophische Untersuchungen*, Suhrkamp, Francfort, 1984, n. 659.

Los aspectos mentales y físicos que la actividad humana presenta quedan conectados entre sí, en la medida en que se explica lo que el agente ha realizado, esto es, el porqué lo ha hecho y el fin que perseguía con su actuación. Son las *razones* que el individuo aporta las que dan unidad y sentido a su actuación. En última instancia, el motor de nuestras acciones se encuentra en las intenciones, no en los estímulos externos. De hecho, es difícil comprender una acción (por ejemplo, al hombre que se le cae un libro y lo recoge) si atendemos únicamente a los estímulos externos.

Al observar la acción humana, ya sea ésta individual o social, surgen bien a las claras sus características específicas, lo cual impide ofrecer para ella un tipo de explicación científica al estilo del modelo de cobertura legal de C. G. Hempel.[9] Es decir, la acción humana no cabe en los estrechos moldes de un tipo de explicación afín al que ha dominado buena parte del panorama de la Física, pues un marco explicativo como el nomológico-deductivo dejaría fuera elementos importantes de la acción humana, justamente aquellos que están más relacionados con el plano de la intención.

Para R. Tuomela, la intención es, junto a las creencias y el querer, uno de los elementos fundamentales en toda acción, ya que se encuentran estrechamente ligados al sentido y finalidad de la acción.[10] Estos elementos tienen una estructura teleológica. Sin embargo, para entender la acción humana, Tuomela no renuncia al concepto de «causa», ya que —a su juicio— hay estados disposicionales que, debidamente actualizados, «causan el comportamiento, o por lo menos, se dan en una constelación de estados que causan el comportamiento».[11]

Desde una perspectiva básica, la intención es un componente crucial en las acciones que realizamos. Las acciones humanas son, en principio, acciones intencionales, puesto que tienen su origen en la intención formada por el agente. La intención es anterior a la propia acción. Supone la existencia de una actividad consciente en el agente. Más aún, es el elemento que da sentido a nuestras acciones y, de alguna manera, también las guía. Se combina con otros elementos como los deseos, los motivos, las obligaciones o normas de cualquier índole.

9. Cfr. HEMPEL, C. G., *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, Free Press, N. York, 1965.

10. Cfr. TUOMELA, R., *A Theory of Social Action*, Reidel, Dordrecht, 1984, p. 80.

11. TUOMELA, R., *Ibidem*, p. 81.

Tal vez sean los *motivos* lo más difícilmente diferenciable de las intenciones. Se podría considerar que son los motivos lo que mueven a obrar (el motivo es el impulso para la acción). Así, al matar a alguien, puede el agente decir que lo hizo «para deshacerse del muy miserable», o que «se lo tenía merecido», o algo similar. Como señala G. E. M. Anscombe, con ello no expresa el agente su intención, sino el motivo de su acción. Porque su intención es claramente quitarle la vida. «La intención de un hombre es a lo que tiende o elige; su motivo es lo que determina su tendencia o elección.»[12] Los motivos explican las acciones, no las originan.[13]

El motivo es previo a la intención, puede ayudar a formar determinadas intenciones. Es este un elemento clave en la formación de las intenciones para actuar, de modo que explicitar el motivo de una acción es verla desde una determinada perspectiva. En algunas ocasiones el motivo y la causa pueden coincidir. En estos casos, el motivo no implica ningún elemento discursivo, sino un hecho pasado (y conocido por el agente) que influye decisivamente en su acción. El motivo, al igual que las causas, mueve a obrar, pero no lo hace así necesariamente. Un sujeto puede tener muchos motivos para actuar y, sin embargo, abstenerse. Los motivos también pueden ser externos al sujeto, como es el caso de la «presión normativa» del ambiente, que de modo externo intenta moverle a actuar.

Quien más ha profundizado en esta cuestión de la «intención» ha sido G. E. M. Anscombe. A su juicio,[14] la *intención* es un acto mental previo a la acción. Comprender una acción —una actuación concreta— supone conocer la intención a la que está ligada; consiste en conocer la intención situada en el origen de la acción. Ve entonces la intención como un acto mental necesario, ya que toda acción intencional se realiza con una intención previa. Sin embargo, este acto mental no se diluye en la acción: la *acción* es la ejecución material de un acto mental. Es la intención la que da sentido a la acción. Así, comprender una acción es entonces comprender la intención que la ha originado.

Dentro de esta perspectiva, comprender una acción comporta conocer la descripción bajo la cual esa acción es intencional. Comprender una acción es ser capaz de aportar los motivos, sentimien-

12. ANSCOMBE, G. E. M., *Intention*, B. Blackwell, Oxford, 1957 (reimpresión de 1976), p. 18.

13. De hecho, es la intención la que origina la acción.

14. Cfr. ANSCOMBE, G. E. M., *Ibidem*; y ANSCOMBE, G. E. M., «Intention», *Proceedings of the Aristotelian Society*, v. 57, (1957), compilado en ANSCOMBE, G. E. M., *Metaphysics and the Philosophy of Mind*, B. Blackwell, Oxford, 1981, pp. 75-82.

tos, intenciones [...] con los que el agente realiza su acción. Requiere, fundamentalmente, poder captar las razones que hacen inteligible una acción. Para Anscombe, la intención es algo «cuya existencia está puramente en el plano de la mente; y, a pesar de eso, [...] surge en la acción».[15] Atañe tanto a lo mental —lo no observable— como a las acciones —lo observable—, las cuales pasan por este motivo a ser denominadas «acciones intencionales».

La principal repercusión de este enfoque para las Ciencias Humanas y Sociales estriba en que su *objeto* de estudio se vuelve más vulnerable, porque queda unido al plano subjetual: las acciones quedan íntimamente ligadas a la intención que las ha producido. Las acciones no se podrían describir correctamente si se hace abstracción de estos elementos mentales. Acaso quedaría patente una estructura radicalmente distinta de la que tienen los objetos de las Ciencias de la Naturaleza, puesto que el ámbito de la intención de la acción es el campo donde en parte se dirime la libertad humana.

Para von Wright, la comprensión entra dentro del ámbito de las razones y motivaciones humanas. «La comprensión no es un tipo de sabiduría. Comprender no implica exactitud. Podemos comprender algo como una acción y, al mismo tiempo, malinterpretarla, esto es, no unirla —como decimos— con las verdaderas razones.»[16] Tuomela muestra, a este respecto, una sintonía, en cuanto que ambos resaltan el papel de la «comprensión». La acción humana posee una consistencia distinta de la que tiene el mundo material existente a nuestro alrededor. Su existencia es, en cierto modo, más precaria; al menos si se la compara con objetos físicos. En otras palabras, ¿se puede hablar de «lo social» como algo real en sentido diferente a otras realidades (como, por ejemplo, la propia de estas páginas)? Los objetos que estudian las disciplinas sobre el hombre y la sociedad no son siempre directamente observables.

Tanto la acción individual como la acción social, se pueden plantear como objeto de estudio en las Ciencias Humanas y Sociales. Según Tuomela, la acción social es «una acción realizada por varios agentes, los cuales —gracias a un acuerdo explícito o implícito— relacionan sus acciones individuales con las acciones de los otros, para lograr una meta conjunta mediante el seguimiento de algunas reglas, prácticas o cosas comunes».[17] En su libro *The Impor-*

15. ANSCOMBE, G. E. M., *Intention*, p. 9.
16. WRIGHT, G. H. VON y MEGGLE, G., «Das Verstehen von Handlungen», *Rechtstheorie*, v. 20, (1988), p. 13.
17. TUOMELA, R., «The Social Dimension of Action Theory», *Daimon*, v. 3, (1991), p. 146.

tance of Us,[18] Tuomela ofrece una caracterización de la acción social centrada en aspectos que resultan básicos en cualquier investigación de tipo psicológico o sociológico. Es un enfoque amplio, pero tal vez no llegue a constituir una teoría explicativa de la acción, sino simplemente sea la *descripción* de los elementos implicados en ella.

Para Tuomela, una acción social es intencional si hay planes conductores para cada uno de los agentes que participan en la realización, de modo que estos agentes ejecutan la acción a propósito, gracias a su plan conductor social, donde cada plan individual posee una referencia esencial a una intención común compartida por los agentes.[19] Esa intención común es ontológicamente una intención individual, que implica una meta y la creencia compartida (esto es, que los demás agentes comparten la misma intención: *we-intention*).

Así pues, al tratar del *objeto* de las Ciencias Humanas y Sociales se aprecia que estamos ante entidades que poseen unas características específicas. Los objetos de estas Ciencias se estudian en su verdadera dimensión al resaltar la «intención», habitualmente asociada a la comprensión (*Verstehen*) y generalmente minusvalorada por los partidarios de la «explicación» (*Erklären*). Hay factores que son propios de los saberes sobre el hombre y la sociedad, aspectos que han de ser «comprendidos» y no meramente «explicados».

Al tener en cuenta la intención y no la simple conducta, se puede llegar al sujeto de las mismas (a su mundo interno, en la medida en que lo manifieste en acciones), pues la actividad desplegada revierte al agente que la despliega. Así se alcanzan planos más profundos, sin permanecer —como tantas veces ocurre— en lo inmediato: en la conducta. Al evaluar desde un marco teórico más amplio, cómo es intención y acción, se puede buscar los elementos previos a la actuación misma, aquellos que se sitúan en el plano de lo mental. Tales aspectos tienen que ver con la acción. Hace falta tener en cuenta el sentido y la finalidad del actuar, sabiendo que ese actuar difícilmente será idéntico en casos posteriores.

Esta condición comporta una dificultad para la experimentación, que es el tipo de contrastación preferido en Ciencias de la Naturaleza. Entendida en sentido estricto, resulta difícil de aplicar en Ciencias Humanas y Sociales, pues las acciones humanas pocas ve-

18. Cfr. TUOMELA, R., *The Importance of Us*, Stanford University Press, Stanford, 1995.

19. TUOMELA, R., *A Theory of Social Action*, p. 154.

ces son reproducibles, con la consiguiente merma para su predecibilidad. Así, en los estudios sobre el hombre y la sociedad, la mirada no ha de estar dirigida hacia leyes más o menos invariantes de las acciones humanas individuales y sociales,[20] sino que primeramente se debe hacer inteligible la actividad humana, indagando su sentido o su finalidad.

Hay, además, una dificultad añadida, que es más clara en Ciencias Humanas y Sociales que en los saberes acerca de la Naturaleza: la necesidad de *interpretar*, que suele remitir al plano de la intencionalidad. Esto sitúa a las Ciencias Humanas y Sociales en unos ejes distintos de la mera «explicación» (*Erklären*): no bastan las medidas cuidadosas de fenómenos reproducibles o predecibles (frecuentemente con un alto nivel de precisión) sino que, para que sean plenamente inteligibles, hace falta aclarar el sentido de las acciones humanas dentro del contexto de valores que las orientan.[21] Esto no merma su objetividad, pero sí acentúa su especificidad.

3. «Causas» y «razones»

Las reflexiones anteriores no suponen la defensa a ultranza de la comprensión, pues cabe una estrecha relación del binomio «causas»-«razones». Tuomela se plantea la cuestión de si es posible hablar de «causas» en la actuación humana. A este respecto, al preguntar acerca del origen de una acción, ¿se busca una causa? Si se recurre al lenguaje cotidiano, parece que, para aclarar la acción humana, se usan tanto las «razones» como las «causas». Cuando se pregunta a unos niños por qué en vez de jugar están rompiendo los juguetes, éstos pueden responder ofreciendo las razones de su propia actuación: «estábamos jugando». Así se obtiene, sin ninguna duda, la descripción de una ejecución material presentada bajo el prisma de la intencionalidad. Se trata de una acción intencional, una acción hecha con una intención: jugar.

No obstante, parece seguro que este estado de la cuestión no va a satisfacer a los padres: ¿por qué actúan de este modo? Ahora parece llegado el momento de las causas. Hay que buscar la causa ori-

20. Una de las más recientes aportaciones a este respecto está en HAUSMAN, D., «¿Necesita leyes la Economía?», *Argumentos de Razón técnica*, v. 3, (2000), pp. 115-138.

21. Cfr. GONZÁLEZ, W. J., «La Ciencia y los problemas metodológicos», en GONZÁLEZ, W. J. (ed.), *Aspectos Metodológicos de la Investigación Científica*, 2.ª ed., Ediciones Universidad Autónoma de Madrid y Publicaciones de la Universidad de Murcia, Madrid-Murcia, 1990, pp. 26-27.

ginante de tal comportamiento, una causa que rebase las explicaciones ofrecidas por el agente acerca de su propia acción. La *causa* se sitúa así en el origen de la acción misma: que hayan visto en una película un comportamiento similar, que esta actuación sea el resultado de algún tipo de malos tratos recibidos, o cualquier otra que se encuentre tras un análisis detenido de su situación y entorno familiar.

Cabe admitir la existencia de ambas —«razones» y «causas»— en torno a la acción intencional. Las *causas* se presentan como un cierto condicionante de la acción. En cambio, las *razones* que da el agente de su actuación pueden o no incorporar las causas de su actuación. En la medida en que sí las recoja tendrá un dominio más pleno de su propia actuación. Al ser preguntado sobre su acción, el agente responde con lo que se puede calificar como «las razones de su actuar». Estas razones pueden también recoger alguna causa.

Las causas pueden ser diversas. En cualquier caso se trata de algo que nos ha movido a obrar, que se produce con anterioridad a las razones y que dirigen la acción en un sentido o en otro. Las causas, a diferencia de los motivos, son elementos ajenos a la voluntad del agente y de los que no siempre se tiene conciencia inmediata.

Conviene resaltar que las «causas» de las que se está tratando aquí no se presentan del mismo modo que las causas en Ciencias de la Naturaleza. Las causas en las Ciencias Humanas y Sociales no determinan el efecto: puesta una causa no podemos afirmar que se sigue el efecto; las causas no determinan la acción, esto es, no presentan una relación de necesidad con respecto a las acciones (que, en tal caso, serían sus efectos), pero nos permiten analizar las acciones atendiendo a determinados patrones explicativos. Frente a esto, en el dominio de la Naturaleza, la causa conlleva el efecto.

En la explicación de la acción humana, no es posible identificar las «razones» con las «causas» sin alterar la especificidad de las Ciencias Humanas y Sociales. Las «razones» conllevan un elemento discursivo acerca de la propia acción que se presenta como irrenunciable para este tipo de Ciencias. Las razones envuelven al *objeto* de estas Ciencias en un plano conceptual, quedando éste caracterizado por dos dimensiones: una material, la ejecución conocida por observación; y otra conceptual, expresable mediante razones que sitúan la acción bajo el prisma de la intencionalidad, lo que implica la irreductibilidad de la acción a lo puramente material.

Parece, pues, que caben ambas —razones y causas— al hablar de las acciones humanas. De hecho, en vez de marcar una frontera tajante entre el mundo de la Naturaleza y el ámbito de las *Geistes-*

wissenschaften, Tuomela busca, dentro del campo de lo humano y social, razones en las acciones intencionales, pero también causas. El patrón explicativo que usa para las acciones intencionales atiende tanto a la razón de la actuación como a sus causas.

4. Teoría *Purposive-Causal*

Mediante lo que se conoce como la teoría «propositivo-causal»,[22] Tuomela explica la estructura de la acción humana. Esta teoría intenta abrir un *nuevo camino* entre las explicaciones teleológicas acerca de la naturaleza humana y las explicaciones meramente causales. Se presenta como un intento de recuperar la «explicación» (en el sentido de *Erklären*): supone, de hecho, la relación de la acción respecto a las «causas» dentro del ámbito de las Ciencias Humanas y Sociales.

Este pensador califica su propia teoría como una explicación de la acción de tipo «teleológico-intencional». Establece una correlación entre la acción que tiene un propósito y la explicación teleológica. «De acuerdo con la Teoría propositivo-causal, una acción humana puede ser considerada como una secuencia de eventos que incluyen unos componentes necesarios: (1) un evento que suscita una actitud propositiva [esto es, orientada a fines], que activa motivacionalmente; (2) un comportamiento corporal, y (3) un resultado público. En el caso de la acción intencional, la actitud propositiva en cuestión es el intentar (realizado o concurrente) del agente».[23]

Tuomela entiende por «acción» un quehacer de carácter público: una secuencia de eventos que incluye elementos observables (una determinada conducta, un movimiento o su omisión) y también otros aspectos no observables: factores motivacionales. Tuomela sitúa en el mismo plano tres bloques globalizadores: 1) el campo de la intención; 2) el ámbito del querer, y 3) el plano de las creencias. Estos tres elementos son lo que denomina «estados disposicionales».[24] A pesar de tener una estructura teleológica, poseen un cierto carácter causal: intenta así conciliar causalidad con finalidad. La causa no ha de entenderse en el sentido de algo que «determina», sino más bien de aquello que «origina». La intención queda así enraizada en los deseos del agente.

22. Cfr. TUOMELA, R., *A Theory of Social Action*, pp. 79-110.
23. *Ibidem*, p. 84.
24. Cfr. TUOMELA, R., *Ibidem*, p. 80.

Introducir en el análisis de la acción el elemento causal no implica que Tuomela reduzca el concepto de «intencionalidad» a la noción de «causalidad». De hacerlo así, nos encontraríamos ante un grupo de acciones que no podemos explicar: los encadenamientos de acciones fortuitas. Veamos un ejemplo que expone el propio autor en *A Theory of Social Action*. «(i) Un cierto personaje quiere matar a su tío; (ii) cree que si atropella a su tío lo matará; (iii) este deseo y esta creencia lo excitan de tal modo que, accidentalmente, atropella a un peatón, y éste, sin saberlo él, no resulta ser otro que su propio tío».[25]

Se trata de un ejemplo de lo que se ha denominado «cadenas causales fortuitas de acciones» (*wayward causal chains*). En rigor, no se puede decir que nuestro agente matase a su tío intencionalmente, no sólo por el hecho contingente de no identificar en aquel instante al peatón con su tío, sino también por la cuestión teórica de que la acción intencional lo es siempre *bajo una determinada descripción.*[26]

Este tipo de ejemplos mueven a Tuomela a formular una teoría donde se pueda contar con las «causas» como elemento explicativo de acciones, pero sin que la intencionalidad quede reducida a términos meramente causales. La intencionalidad se mantiene como elemento plenamente teleológico, y las acciones se explican en función de un componente propositivo, que implica un plan conductor, esto es, una finalidad.

Según el enfoque de Tuomela, dentro de la estructura teleológica de la acción tanto la «intención» como la «creencia» son conceptos mentales que poseen gran importancia explicativa. Ontológicamente, son anteriores al comportamiento, ya que están en su origen y, en cierto sentido, lo causan. Las intenciones quedan unidas a un querer. La acción humana incluye, por tanto, la intención y el querer. Estas nociones no son sinónimas, sino que, el agente forma intenciones para actuar, cuando las creencias y los deseos internos se activan.

El *querer* de un agente se relaciona con la acción que, en cierto sentido, «causa». Entre la acción y el querer hay un doble nexo:

25. *Ibidem*, p. 61.

26. Las razones implican un elemento discursivo que forma parte de la acción misma. Ésta se caracteriza por dos dimensiones: una material, la realización efectiva que es conocible mediante observación; y otra conceptual: expresable mediante razones que sitúan la acción bajo el prisma de la intencionalidad. Esto implica que la acción es irreductible a lo puramente material y que sólo cuando se aportan las razones de una determinada acción se proporciona la descripción bajo la cual ésta es intencional.

el querer actúa en un orden causal, tanto al dar lugar a la acción como en relación a los resultados de dicha acción. Según Tuomela, en el caso de no ocurrir el resultado deseado, cabe decir que se trata de un mero intento o de un intento sin éxito, pero no se puede saber si el agente quería realmente conseguir un determinado resultado o no. El querer es un evento consciente que actúa y origina una determinada intención: tiene la función de dirigir y controlar la acción. En esta línea, la acción realizada la define en los siguientes términos: «una acción realizada tiene la estructura de una secuencia finita de eventos "v,... b,... r". En esta secuencia de eventos, "v" representa una causa mental del movimiento corporal público "b". El evento resultante "r" representa la realización de la acción».[27]

La noción técnica de *plan conductor* es otro elemento a destacar en la teoría propositivo-causal. Se trata del plan del agente para actuar. El plan conductor de un agente singular es teóricamente «una generalización del concepto de silogismo práctico, o mejor, de las premisas de un silogismo práctico».[28] Tuomela defiende la independencia lógica entre las premisas y la inferencia práctica; además, esta noción no implica necesariamente que la acción haya sido planeada, estructurada con más o menos exactitud, ya que entonces no tendrían cabida las acciones habituales o rutinarias dentro de la noción de «acción intencional».

Sin embargo, el plan conductor también puede ser inconsciente. Este caso se da cuando hay acciones habituales, casi automáticas o aparentemente irreflexivas, de modo que las acciones surgen como respuesta a una demanda. Este tipo de acciones son intencionales, puesto que, en cualquier momento, el agente —al ser preguntado— puede dar razón de su actuación. Las acciones habituales de las que quisiéramos afirmar que son plenamente intencionales, lo son también en este esquema.

Ahora bien, el plan conductor y la acción se sitúan a niveles diferentes. Las intenciones dirigidas hacia el futuro surgen al valorar el plan conductor, mientras que las intenciones-en-la-acción se dan en el plano de la realización material, de la acción misma. Teniendo en cuenta este enfoque, Tuomela explica la acción intencional en la teoría propositivo-causal del siguiente modo: «(PC) un agente realizó la acción «U» de modo intencional si y sólo si

27. TUOMELA, R., *Human Action and Its Explanation*, p. 290.
28. *Ibidem*, pp. 85-86.

1) el comportamiento del agente «b» genera el resultado «r» a propósito;
2) había un plan conductor, «k», del agente que comporta un fin, que el agente intenta realizar mediante su comportamiento corporal (del tipo que cree ejemplificado por «b»); y cree que con su comportamiento lo llevará a cabo o, al menos, le conducirá hacia él; y
3) este intentar efectivo (como un querer, «T»), juntamente con su creencia, generan a propósito el comportamiento «b» del agente «U».[29]

Este concepto plantea algunos inconvenientes en la *acción social*. En éstas debe existir una meta común que afecta a todos los agentes. Esta meta o fin común cohesiona la acción conjunta; implica un cierto acuerdo. Es el acuerdo lo que constituye las acciones individuales en una verdadera acción conjunta, ya que es en el acuerdo donde se decide conjuntamente hacer algo de modo común. En la acción social, los planes de conducta se explican como «conjunciones especiales de los planes de conducta individuales de los agentes que participan».[30] El plan conductor de una acción es algo más que la exposición mental de una intención; recoge las premisas de un silogismo práctico: «A» quiere «q», a no ser que haga «p» no logrará «q».

Tuomela mantiene que no existen diferencias estructurales básicas entre la acción individual y la acción social; pero la conexión entre los agentes debe quedar suficientemente clara. Esto no presenta problemas en acciones realizadas por pocos agentes (por ejemplo, dos personas en el momento de casarse o cuatro jugando a las cartas). En las realizadas por muchos individuos, se pueden plantear otras circunstancias, sobre todo si la acción es multitudinaria. En éstas no se exige que el plan conductor sea un elemento consciente y conocido en todos sus términos por todos los agentes participantes; no obstante, sí es exigible que, por lo menos, esté presente en un número suficientemente representativo de sujetos.

Tomando una manifestación como ejemplo de acción de este segundo tipo, la actuación de aquellos individuos que no mantienen ya la intención que les movió a obrar se dejaría de considerar como parte de la acción social. Podría darse el caso paradójico de que es-

29. TUOMELA, R., «Intentional, Single, and Joint Action», *Philosophical Studies*, v. 62, (1991), p. 249.
30. TUOMELA, R., *A Theory of Social Action*, p. 86.

tos agentes —que ya no mantienen la intención inicial— sigan estando presentes en el grupo. Al haber perdido el fin de su propia acción, ya no se puede decir que éstos están manifestándose: simplemente deambulan por la calle con los demás.

Según Tuomela, «los planes de conducta son necesarios en dos lugares centrales en la Teoría propositivo-causal de la acción. En primer lugar, se necesitan en la elucidación del concepto de acción intencional. En segundo término, juegan un papel central en la teoría sobre la explicación de la acción social».[31] Participan en la explicación de la «acción intencional», en la medida en que recogen las premisas del silogismo práctico. Implican, por tanto, la especificación de los medios a poner para obtener una meta propuesta. En la acción social este elemento forma parte del acuerdo que la estructura.

Así pues, los rasgos más característicos de esta teoría de Tuomela son dos: *a*) la causalidad exigida es propositiva (esto es, está orientada a fines), y *b*) la acción intencional se entiende ontológicamente como una secuencia ordenada: cuenta con intenciones, creencias y deseos relevantes; y no se reduce simplemente a una intención, sin más. No mantiene que el *querer* «cause» literalmente la acción, sino que se trata, en parte, de una noción causal.

5. Explicación de la acción intencional

En general, la teoría expuesta ofrece —a mi juicio— un análisis de la acción intencional acorde con la estructura que la misma parece tener en realidad. Intenta buscar una vía intermedia entre la explicación causalista, propia de *Erklären*, y la teleológica, característica de *Verstehen*. A pesar de ello, el mismo Tuomela señala que una interpretación de esta teoría que intente situarla entre las explicaciones plenamente causales y las teleológicas desfigura su exposición. Considera que entre el «querer» y la «intención» no existe una conexión causal, a pesar de que la lectura de sus textos así lo induce a pensar.[32]

De sus escritos, sin embargo, sí parece inferirse que es el *querer* y su relación con la intención el lugar donde se instala el elemento causal. En mi opinión, es en el paso del querer (esto es, del elemento volitivo) a la intención misma donde Tuomela introduce el elemento causal, dentro de la cadena implicada por una acción

31. *Ibidem*, p. 86.
32. Cfr. TUOMELA, R., *Ibidem*, pp. 79-93.

intencional. La *intención* no puede ser reducida a puro querer y, además, el querer genera causalmente la intención. En esta línea, señala expresamente que «las intenciones son concomitantes a la acción. Aunque son irreductibles al querer, se puede considerar que exigen la presencia del querer, ya sea éste un querer intrínseco (algo deseado por sí mismo) o extrínseco (por ejemplo, las obligaciones). [...] Se puede decir que, tanto el deseo intrínseco (junto con otras determinaciones efectivas internas) como el deseo extrínseco acompañado por las creencias apropiadas (determinaciones doxásticas), pueden generar —y de hecho generan, al ser conjugadas convenientemente— la intención de modo causal».[33]

Hay en el individuo estados disposicionales que manifiestan la actitud del individuo. Estas *disposiciones* se introducen como causas potenciales del comportamiento. Para que lo sean realmente, deben causar la intención: han de presentarse en una circunstancia adecuada como estados activantes. Las actitudes propositivas activan una disposición causando el comportamiento; contribuyen al surgir de la acción, porque son la causa de la intención situada en el origen de la ejecución material o elemento externo de la acción. En la medida en que se dice del querer que *activa* una disposición, es posible mantener que éste es el propio flujo de la conciencia que ha de causar la intención. El querer, al igual que la creencia, dependen en su origen de algún tipo de estímulos, esto es, son estímulo-dependientes y, además, son proposicionales. Tuomela no defiende ningún tipo de determinismo psicológico, ya que la acción misma a la que apunta la intención no queda implicada de modo necesario.

Conviene resaltar que Tuomela no defiende la causalidad entendida al modo de las Ciencias de la Naturaleza. En este sentido, se podría aplicar aquí la distinción de von Wright —ya mencionada— entre «origen» y «causa».[34] Para Tuomela, el querer y la intención no quedan identificados. Cuando un agente intenta algo, cree poder lograrlo. El ser humano puede, sin embargo, desear muy vivamente cosas que le resulten imposibles de conseguir. Al deslindar el querer de la intención, surge la cuestión de indagar qué relación guarda el querer con la acción. La respuesta de Tuomela es clara: la meta. El contenido de un querer es la meta del agente. El querer es lo que le mueve a actuar. El querer motiva al agente a realizar acciones que, a su juicio, son adecuadas para lograr una determinada

33. TUOMELA, R., «Intentional, Single, and Joint Action», p. 235.
34. Cfr. WRIGHT, G. H. VON, *Explanation and Understanding*, pp. 132 y 135.

meta. El querer genera, por tanto, una intención cuya descripción queda recogida en el plan conductor. Con respecto al fin, el querer se diferencia de la intención, en cuanto que el objeto del primero es un estado externo al individuo, mientras que la intención tiene por objeto nuestras propias acciones.

Se dice de una acción social o individual que responde a razones si, ante la pregunta acerca de por qué ha sido realizada una determinada acción, se obtiene una respuesta que hace depender la acción entera de un razonamiento práctico. En cambio, se sostiene que la unión entre los antecedentes de la acción y la ejecución de la acción es debido a causas cuando se entiende que la acción realizada «por un agente singular se puede construir como un proceso causal que se inicia con una volición [...], a la que se añade como componentes suyos el movimiento corporal (generado por la volición y el resultado externo producido)».[35]

Mantiene así claramente que la *acción* es, en sí misma considerada, un proceso causal: «la conducta es un proceso causal que se inicia con una causa interna y culmina en un comportamiento público».[36] Tuomela conjuga este planteamiento con la exigencia de explicar la acción (en el proceso causal citado) a tenor de *razones* que expliciten la intención que hay tras ella. Es entonces cuando opta por la explicación teleológica.

Debido a este uso de la noción de «causa», resulta sugerente la lectura de la explicación de la acción social ofrecida por Tuomela, porque parece conllevar una aproximación de las explicaciones causales a las explicaciones teleológicas. De alguna manera, cabe mantener que los antecedentes de la acción incluyen la existencia de una causalidad por fines (*purposive causation*). Sin embargo, esta posible convergencia no es total, ya que lo que tenemos es un uso laxo del término «causa», empleado en el sentido de «origen». Tuomela reconduce en su modelo explicativo de las acciones humanas la intención como rasgo central.

La teoría *purposive-causal* pretende mostrar la estructura interna de la acción intencional reivindicando la existencia de «causas», sin usar este término en su sentido más restrictivo. Se diferencia de otras propuestas de tipo causalista en que no reduce la intención a esquemas causales al exponer la estructura de la acción. En su postura, los antecedentes de la acción son de tres tipos: el

35. TUOMELA, R., «Are Reason-explanations Explanations by Means of Structuring Causes?», *Philosophy and Phenomenological Research*, v. 1, n. 4, (1990), p. 813.
36. *Ibidem*, p. 814.

querer, la creencia y la intención. Entre «querer» e «intención» se da una relación causal: el querer es causa de la intención. Pero no se debe entender *causa* en el sentido de «determinar en efecto», sino como ya se ha señalado: en cuanto que es su origen. Concibe la creencia como un componente que nos hace juzgar acerca de las oportunidades para la acción que posee el individuo, y ve la creencia en la acción social como decisiva en su estructura, pues la creencia contribuye a la intersubjetividad conceptual que le es necesaria a la acción social. Se está hablando de la creencia común de que todos los agentes actúan con la misma intención.

En lo desarrollado hasta el momento, se observa cómo Tuomela sitúa en el origen de la acción tanto razones como causas. El agente ofrece razones para presentar su acción como una actividad inteligible, que persigue alguna meta determinada. Pero, él estructura en términos causalistas las relaciones entre el conjunto de los antecedentes que ofrece la acción y la acción propiamente dicha. A este respecto, mantiene que el querer causa la intención. Ahora bien, no llega a adoptar un enfoque característico del *Erklären* porque el término «causa» no es entendido en el mismo sentido que en las Ciencias de la Naturaleza. En otras palabras, no lo interpreta como «determina», sino más bien como «origina».

6. **Consideraciones finales**

Con frecuencia, la teoría de Tuomela describe bastante bien lo que son realmente las acciones humanas. Así, nuestras acciones, además de ser productos libres, de los cuales los agentes pueden dar razón, responden en otros aspectos a «elementos» de los que no podemos dar razón, y que —en algunas ocasiones— son relativamente ajenos al propio agente. Es en este momento donde quisiéramos hablar de «causas». De hecho, el concepto de *causa* surge en las Ciencias Humanas y Sociales, aunque, obviamente, no sean lo mismo que causas de la Naturaleza. Los eventos a los que podríamos denominar «causas» forman una parte más del desarrollo natural del individuo. Son elementos que realmente explican las acciones humanas y, sin embargo, no recogen las razones que movieron al individuo a obrar. A veces estos elementos ya no están presentes a la hora de actuar, pero han marcado al ser humano lo suficiente como para poder ser eventos que realmente influyen en nuestras acciones.

Esto conecta plenamente con la experiencia particular de cada individuo: nuestras actuaciones son, en mayor o menor grado, re-

gulares; la vida del ser humano se compone primordialmente de acciones, muchas de las cuales reitera. Surge aquí una cierta regularidad que parece deberse más a causas que a las razones que el agente aduce; es más, incluso se puede llegar a sostener que, en determinadas circunstancias, las acciones de las personas son «bastante predecibles». Es por ello que la situación se presta a hablar de «causas», cuyo conocimiento junto a las razones del propio sujeto, aclaran las propias actuaciones.

Como teoría explicativa concreta de la acción intencional, Tuomela propone la teoría «propositivo-causal». Se diferencia de otras propuestas de tipo causalista en que no reduce la intención a causalidad cuando explica la estructura de la acción. En su postura, los antecedentes de la acción son de tres tipos: el querer, la creencia y la intención. Entre «querer» e «intención» se da una relación causal: el querer es causa de la intención. Pero no se debe entender causa en el sentido de «determinar un efecto», sino en cuanto que es su origen. La creencia se concibe como un componente que nos hace juzgar acerca de las oportunidades para la acción que posee el individuo; en la acción social es un elemento decisivo, pues la creencia contribuye a la intersubjetividad conceptual que le es necesaria a la acción social.

Básicamente, el planteamiento de Tuomela está en consonancia con lo que, de hecho, ocurre en la investigación en las Ciencias Humanas y Sociales. Hay así elementos, como las causas, que aclaran nuestras acciones; sin embargo, no son meras razones, ni motivos que guíen acciones. A veces, estos elementos ya no están presentes a la hora de actuar, pero han marcado al ser humano lo suficiente como para poder ser eventos que realmente influyen en nuestras acciones. Muchas veces el individuo no es consciente de esta situación.

Esta propuesta de acortar distancias entre las Ciencias Humanas y Sociales y las Ciencias del ámbito de la Naturaleza —en la medida en que se incorporan causas en la explicación de las acciones humanas— no implica caer en un determinismo; no se renuncia a la libertad, que es uno de los rasgos distintivos del ser humano. El papel de la intención (el elemento discursivo inicial) sigue siendo crucial para la actuación. Éste es el elemento más destacado dentro de los antecedentes de la acción. La intención se sitúa en el origen de la acción individual, al implicar la meta que se pretende y ejercer sobre la realización de la acción un poder de control.

Al buscar en el origen de la acción causas y razones, se introduce en el modelo teórico aspectos que implican la acción de un

modo casi automático, propiciando así una explicación causal que junto a los elementos teleológicos —que ayudan a comprender desde los fines—, aclaran la acción. Así, el sistema explicativo de la acción es completo, dado que abarca todo tipo de elementos, y sin embargo, la acción humana se sigue considerando como el resultado de desplegar la actividad del sujeto desde los actos mentales que le dan origen.

Sólo se puede hablar de causas en las acciones humanas cuando éstas han sido efectivamente realizadas. Las causas que encontramos en el origen de la acción no implican la acción de modo necesario, simplemente la presuponen con un cierto grado de probabilidad. A pesar de ello —y, sobre todo, cuando este conocimiento teórico acerca de la acción pretende ser práctico— no son reducibles a las razones que el agente ofrece de su actuación, ya que muchas veces escapan a su dominio. Las Ciencias Humanas y Sociales integran, en la teoría explicativa de la acción humana, tanto razones —elemento discursivo— como causas, elemento externo y ajeno al agente, que en muchas situaciones permanece en el campo de lo desconocido o no dominable.

Referencias bibliográficas

ANSCOMBE, G. E. M., *Intention*, B. Blackwell, Oxford, 1957, reimpr. 1976.

DILTHEY, W., *Einleitung in die Geisteswissenschaften. Versuch einer Grundlegung für das Studiem der Gessellschaft und der Geschichte*, en DILTHEY, W., *Gesammelte Schriften*, Teubner y Vandenhoeck-Ruprecht, Stuttgart y Gotinga, 5.ª edic., 1962.

ELSTER, J., *Solomonic Judgements*, Cambridge University Press, Cambridge, 1989. Vers. cast. de C. Gardini: *Juicios Salomónicos*, Gedisa, Barcelona, 1991.

GONZÁLEZ, W. J., «La Ciencia y los problemas metodológicos», en GONZÁLEZ, W. J. (ed.), *Aspectos Metodológicos de la Investigación Científica*, 2.ª ed., Ediciones Universidad Autónoma de Madrid y Publicaciones Universidad de Murcia, Madrid-Murcia, 1990, pp. 15-46.

GONZÁLEZ, W. J. (ed.), *Action Theory—Teoría de la Acción*, número monográfico de *Daimon*, v. 3, (1991).

GONZÁLEZ, W. J., «Caracterización del objeto de la Ciencia de la Historia y bases de su configuración metodológica», en GONZÁLEZ, W. J. (ed.), *Acción e Historia. El objeto de la Historia y la Teoría*

de la Acción, Publicaciones Universidad de A Coruña, A Coruña, 1996, pp. 25-111.

GONZÁLEZ, W. J., «From *Erklären-Verstehen* to *Prediction-Understanding*: the Methodological Framework in Economics», en SINTONEN, M., YLIKOSKI, P. y MILLER, K. (eds.), *Realism in Action*, Kluwer, Dordrecht, 2002, pp. 38-48.

HEMPEL, C. G., *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*, Free Press, N. York, 1965.

MANNINEN, J. y TUOMELA, R. (eds.), *Essays on Explanation and Understanding*, Reidel, Dordrecht, 1977.

RESCHER, N., *Razón y valores en la Era científico-tecnológica*, Paidós, Barcelona, 1999.

TUOMELA, R., *Human Action and Its Explanation*, Reidel, Dordrecht, 1977.

TUOMELA, R., *A Theory of Social Action*, Reidel, Dordrecht, 1984.

TUOMELA, R., «Are Reasons-Explanations by Means of Structuring Causes?», *Philosophy and Phenomenological Research*, v. 1, n. 4, (1990), pp. 813-818.

TUOMELA, R., «The Social Dimension of Action Theory», *Daimon*, v. 3, (1991), pp. 145-158.

TUOMELA, R., «Intentional Single and Joint Action», *Philosophical Studies*, v. 62, (1991), pp. 235-262.

TUOMELA, R., *The Importance of Us*, Stanford University Press, Stanford, 1995.

WITTGENSTEIN, L., *The Blue and Brown Books. Preliminary Studies on the Philosophical Investigations*, Basil Blackwell, Oxford, 1978.

WITTGENSTEIN, L, *Philosophische Untersuchungen*, Suhrkamp, Francfort, 1984.

WRIGHT, G. H. VON, «Action and Explanation of Behaviour», *Psykologia*, v. 1, (1966), pp. 132-135.

WRIGHT, G. H. VON, *Explanation and Understanding*, Cornell University Press, Ithaca, 1976.

WRIGHT, G. H. VON y MEGGLE, G., «Das Verstehen von Handlungen», *Rechtstheorie*, v. 20, (1988), pp. 3-34.

CAPÍTULO 9

EXPLICACIONES PROGRAMÁTICAS Y EXPLICACIONES FUNCIONALES

DAVID PINEDA

Los filósofos australianos Frank Jackson y Philip Pettit han defendido en los últimos años la idea de que si bien las propiedades funcionales no son causalmente eficaces son, sin embargo, causalmente pertinentes, en el sentido de que mencionarlas resulta crucial en ciertos contextos explicativos que ellos denominan «explicaciones programáticas». Su propuesta puede resumirse en la tesis central de que la mención de ciertas propiedades puede resultar crucial en cierto tipo de explicaciones causales de un efecto, incluso aunque no sean causalmente responsables de ese efecto. Dicha tesis se plantea como una defensa de la pertinencia causal de las propiedades psicológicas, entendidas como propiedades funcionales, frente a ciertos influyentes argumentos escépticos. En el presente artículo se propondrá un análisis refinado de la noción de «explicación programática», análisis que —a mi juicio— es necesario ofrecer para poder hacer plausible la tesis central de Jackson y Pettit. Sin embargo, como veremos, dado este análisis refinado, su réplica a los argumentos escépticos no va a poder mantenerse. De hecho, una de las cosas que espero mostrar es que la distinción entre propiedades causalmente eficaces y propiedades causalmente pertinentes no es de ninguna ayuda a la hora de responder a esos argumentos. En la última parte del artículo caracterizaré una serie de casos para los que las propiedades funcionales pueden ser vistas como causalmente pertinentes según mi análisis refinado, que considero correcto, de la noción de *explicación programática*.

1. Introducción

Para los fines que persigo en este trabajo, consideraremos una explicación funcional de un efecto E[1] cualquier explicación de E en la cual se menciona crucialmente, o de modo ineliminable, una propiedad funcional. Al decir que la mención de la propiedad funcional es «crucial» o «ineliminable» en una explicación, lo que queremos decir es que esa explicación no sobreviviría la omisión de esa propiedad funcional. Asumiremos también la caracterización usual de una propiedad funcional como aquel tipo de propiedad ejemplificada por algo X cuando y sólo cuando X ejemplifica una propiedad que tiene el rol funcional asociado a la propiedad funcional, donde un rol funcional se especifica siempre a través de un conjunto de relaciones causales.[2] Un realizador de una propiedad funcional será entonces cualquier propiedad que tenga el rol funcional asociado, es decir, que capacite al objeto que la ejemplifique para tener las relaciones causales especificadas. Por ejemplo, digamos que el rol funcional de la propiedad de ser soluble en agua es causar la disolución al ser sumergido en agua; así, de acuerdo con esta caracterización de propiedad funcional, un objeto ejemplifica la propiedad de ser soluble en agua si y sólo si ejemplifica alguna propiedad que causa la disolución en caso de inmersión en agua.

Sea el funcionalismo la tesis de que las propiedades psicológicas, en especial las propiedades intencionales como creer que dentro de poco lloverá o desear comer un helado de chocolate, son propiedades funcionales, esto es, la tesis de que algo X ejemplifica una propiedad psicológica cuando y sólo cuando X ejemplifica alguna propiedad que capacita a X para causar y ser causado en los modos especificados por el rol funcional asociado a la propiedad funcional psicológica.[3]

1. Los objetos de las explicaciones causales, los efectos causales, se entenderán aquí como entidades consistentes en la ejemplificación de una propiedad por parte de un objeto o la ejemplificación de una relación o propiedad relacional por parte de una secuencia de objetos.

2. Esta caracterización sigue el llamado «análisis realista de las disposiciones» (véase Armstrong [1968], Lewis [1986] y Schiffer [1987]) frente al llamado «análisis ryleano» (Ryle [1949]). No es este el lugar donde justificar el análisis realista, que se ha impuesto en las últimas décadas. Para algunos argumentos clásicos en su favor, véase Armstrong (1968), pp. 85-88.

3. El funcionalismo es la teoría predominante hoy en día en Filosofía de la Mente. Entre los trabajos clásicos en los que se propone y argumenta esta doctrina pueden consultarse Putnam (1967), (1973) y también Lewis (1966). Para una exposición clara del funcionalismo y de algunos de sus problemas, véase Block (1978).

Llamemos finalmente «racionalización» a cualquier explicación de una conducta humana que consiste en la atribución al agente de propiedades psicológicas, como creencias o deseos. Esto es, una racionalización no es más que una explicación de sentido común de la conducta de un agente.

Hechas estas precisiones terminológicas, lo que resulta es que si aceptamos el funcionalismo entonces las racionalizaciones devienen explicaciones funcionales. Si aceptamos también la tesis de que las racionalizaciones son un cierto tipo de explicaciones causales, esto es, explicaciones cuyo *explanans* consiste en la mención de causas del *explanandum*, deberíamos también comprometernos con la tesis de que las explicaciones funcionales son explicaciones causales, o al menos con la tesis de que algunas lo son.

Sin embargo, ciertos argumentos que se detallarán en la sección siguiente, parecen mostrar que las propiedades psicológicas en particular y las funcionales en general no son causalmente responsables para la conducta de un agente, esto es, que la causa de que el agente actúe como lo hace no reside en la ejemplificación de esas propiedades. Aceptar esos argumentos, así pues, parece obligarnos a abandonar o bien el funcionalismo o bien la tesis de que las racionalizaciones son explicaciones causales. Pues bien, lo que Jackson y Pettit pretenden argumentar es que este juicio es precipitado porque en realidad los argumentos contra la eficacia causal de las propiedades funcionales y psicológicas son, en contra de las apariencias, perfectamente compatibles con el funcionalismo y la tesis de que las racionalizaciones son explicaciones causales. Argumentan que hay un sentido en el que las explicaciones funcionales son explicaciones causales si bien las propiedades funcionales no son de hecho causalmente eficaces.

Lo que haré a continuación es presentar la propuesta de Jackson y Pettit y argumentar que, entendida correctamente, tal propuesta es incapaz de compatibilizar los argumentos escépticos con las dos tesis referidas. En esto consistirá la parte negativa o crítica del presente trabajo. La parte positiva consistirá en ofrecer un nuevo análisis de la noción crucial de explicación programática, distinto del que ofrecen los propios Jackson y Pettit. El nuevo análisis abrirá las puertas a una teoría acerca de cómo es posible que ciertas propiedades sean causalmente pertinentes en la explicación de un efecto sin ser causalmente responsables del mismo y también a una útil distinción entre explicaciones funcionales informativas y no informativas.

2. La propuesta de Jackson y Pettit

Comencemos por describir brevemente los dos argumentos contra la eficacia causal de las propiedades psicológicas intencionales que Jackson y Pettit tienen en mente. El primero es el argumento del duplicado o «doppelgänger». Este argumento comienza asumiendo el externismo con respecto a las propiedades intencionales. El externismo es la tesis, basada en las consideraciones de Hilary Putnam y Tyler Burge, según la cual las propiedades intencionales se individualizan externamente, es decir, por referencia a objetos y propiedades del entorno del agente que las ejemplifica.[4] Si esto es correcto, entonces parece que las propiedades no sobrevienen sobre propiedades intrínsecas del agente sino que su base de superveniencia debe extenderse también a ciertas propiedades del entorno del agente (y tal vez también a ciertos hechos acerca de la historia del agente).[5] El argumento del duplicado es un argumento contra la eficacia causal de las propiedades externistas. Considérese un duplicado idéntico a mí molécula por molécula pero ubicado en un entorno distinto en el que no haya agua sino una substancia químicamente distinta, llamada «biagua», aunque con la misma apariencia que el agua (es húmeda, es incolora, es inodora, es insípida, etc.). Pues bien, cuando ejemplifico la propiedad de creer que el agua es húmeda, por ejemplo, mi duplicado ejemplifica en realidad otra propiedad intencional, a saber, la de creer que la biagua es húmeda. Ahora bien, al ser bajo suposición mi duplicado, esto es al no haber ninguna diferencia intrínseca, ni siquiera al nivel molecular, entre él y yo, su conducta debe ser idéntica a la mía, al menos en ciertos aspectos relevantes como es el caso de nuestros movimientos corporales. En consecuencia parece seguirse de ello, en aplicación del método de coincidencia de Mill, que las propiedades intencionales, individualizadas externamente, no son causalmente responsables de la conducta (al menos no de los movimientos corporales), y que sólo el llamado «contenido restringido», esto es, esa parte del contenido intencional que sobreviene sobre las propiedades intrínsecas del agente, es decir, esa parte del contenido intencional que comparto con mi duplicado, puede ser causalmente responsable de la conducta.

4. Véase Putnam (1975) y Burge (1979).
5. Sobre la noción de superveniencia, véase Kim (1984); sobre el uso de esta noción en las discusiones acerca de la eficacia causal de la mente, véase Horgan (1993).

El segundo argumento es más general porque concierne a cualquier propiedad psicológica. Si asumimos (como típicamente quiere el funcionalista) que las propiedades psicológicas son múltiplemente realizables[6] por propiedades físicas y asumimos también que puede en principio ofrecerse una explicación causal completa de cualquier movimiento corporal de un agente en términos únicamente de propiedades físicas, se sigue de todo ello que no hay ninguna contribución causal distintiva que las propiedades psicológicas puedan hacer con respecto a esos movimientos corporales. Así, la atribución de responsabilidad causal a las propiedades psicológicas de un agente con respecto a sus movimientos corporales sería en el mejor de los casos redundante, y al carecer de una razón no *ad hoc* para insistir en tal redundancia, lo más razonable parece ser renunciar a tal atribución.[7]

Este argumento puede también hacerse extensivo a cualquier propiedad funcional. Volviendo a nuestro ejemplo, dado que cualquier realizador de la solubilidad en agua causa la disolución en agua (de hecho, es precisamente por ello, como hemos visto, que cabe describirlo como realizador de la solubilidad en agua), parece de nuevo que no hay ninguna contribución causal distintiva que la propiedad funcional pueda hacer con respecto al efecto de la disolución.

La posición de Jackson y Pettit frente a estos argumentos, contraria a la de la mayoría de funcionalistas, es la de aceptarlos, y por consiguiente aceptar también la conclusión de que las propiedades funcionales en general y las psicológicas en particular carecen de eficacia causal. Su esfuerzo se dirige entonces a mostrar que estos argumentos, a pesar de lo que aparentan, no llevan a la conclusión de que esas propiedades no son causalmente pertinentes. Ello es así, de acuerdo con Jackson y Pettit, porque los argumentos son compatibles con el hecho de que la mención de una propiedad funcional sea indispensable en ciertos tipos de explicaciones causales que ellos llaman «explicaciones programáticas», con lo cual no serían después de todo causalmente irrelevantes.

6. No puedo entrar aquí en el análisis de la compleja noción de realización. Para un análisis y discusión interesante de la misma, cabe resaltar a Shoemaker. Sin embargo, para los objetivos de este artículo basta con que el lector suponga que al decir que las propiedades psicológicas son múltiplemente realizables por propiedades físicas todo lo que se quiere decir es que en la base de superveniencia de cada propiedad psicológica aparecen propiedades físicas distintas.

7. Éste es esencialmente el argumento de la exclusión causal propuesto por Jaegwon Kim. Véase Kim (1993) y (1998).

La idea básica de Jackson y Pettit puede resumirse como sigue. Una explicación programática de un efecto E es una explicación en la que se ofrece información no meramente del proceso causal que de hecho ha llevado al efecto E sino también de otros procesos causales contrafácticos que llevarían también, de tener lugar, al efecto E (o, más precisamente, a otro efecto del mismo tipo que E). Ahora bien, esa información causal contrafáctica no puede transmitirse mencionando una propiedad causalmente eficaz para E, sino sólo a través de la mención de una propiedad sobreviniente sobre la propiedad o propiedades causalmente eficaces. De este modo, la propiedad sobreviniente, si bien causalmente impotente con respecto al efecto, resulta indispensable a la hora de obtener cierto tipo de información causal sobre ese efecto, a saber, información causal contrafáctica, con lo cual deviene causalmente pertinente para ese efecto.

Los autores ilustran estas ideas con ayuda de ejemplos en los que la propiedad causalmente eficaz y la propiedad sobreviniente están en la relación de determinado a determinable. Consideremos uno de estos ejemplos. Imaginemos que un ascensor ha sido diseñado de tal modo que se para si suben a él más de diez personas y supóngase que en una determinada ocasión veinte personas suben al ascensor y en consecuencia éste se para.[8] Obsérvese que en una situación de este tipo podemos ofrecer dos explicaciones del hecho de que el ascensor se haya parado. Podemos decir que el ascensor se ha parado porque han subido a él veinte personas o bien que se ha parado porque han subido a él más de diez personas. En un cierto sentido la segunda explicación parece mejor que la primera, incluso aunque aceptemos que la propiedad causalmente eficaz para el paro del ascensor es la propiedad de tener veinte personas a bordo y no la propiedad de tener más de diez personas a bordo. Parece mejor en el sentido de que nos informa no sólo de lo que ocurriría si subieran al ascensor veinte personas, sino también acerca de lo que ocurriría si subieran al ascensor once, doce, o treinta personas. Esta segunda explicación nos da así pues un tipo de información causal contrafáctica acerca del paro del ascensor que no nos ofrece la primera. En este sentido, puede decirse incluso que la segunda explicación nos permite entender mejor qué es lo que ha ocurrido con el ascensor que la primera.

¿Cómo es posible que la mención de una propiedad causalmente impotente para un efecto E proporcione información causal valiosa sobre E, incluso en cierto modo más valiosa que la que proporciona la mención de la propiedad causalmente eficaz? La res-

8. Jackson y Pettit (1990), pp. 205-206.

puesta de Jackson y Pettit es que en el caso de nuestro ejemplo el ascensor ha sido diseñado de tal modo que toda propiedad que pertenece a la base de superveniencia de la propiedad de tener más de diez personas a bordo es causalmente eficaz para el efecto a explicar, esto es, para el hecho de que el ascensor se pare. Ésta es la razón por la que la mención de la propiedad sobreviniente es en cierto modo más informativa acerca del origen causal del efecto, y más apropiada en este contexto explicativo, que la mención de la propiedad subviniente cuya ejemplificación de hecho causa el efecto.

Similarmente ocurre, según Jackson y Pettit, con las propiedades funcionales y sus realizadores: «las explicaciones funcionales [...] nos informan acerca de la clase de estados que producen o producirían el resultado sin informarnos de qué estado en particular produjo de hecho el resultado» (Jackson y Pettit, 1988, p. 396).

Podemos decir por ejemplo que la solubilidad en agua es causalmente pertinente, aunque no causalmente eficaz, con respecto a un caso de disolución por inmersión en agua, porque todo realizador de la solublidad en agua causa la disolución en condiciones de inmersión en agua. Razonando de modo parejo al caso de la propiedad de tener más de diez personas a bordo y la propiedad de tener veinte personas a bordo, podemos decir que en ciertos contextos explicativos (aquellos en los que se persigue obtener información causal contrafáctica acerca del efecto) un caso determinado de disolución tras inmersión en agua se explica mejor y de un modo más adecuado mencionando a la propiedad funcional de ser soluble en agua que mencionando a cualquiera de sus realizadores.

Finalmente, y supuesto lo que se asume en los dos argumentos escépticos anteriores, a saber, que todo realizador físico de una propiedad intencional funcional es causalmente eficaz con respecto a los movimientos corporales, podemos decir que en ciertos contextos la mención de propiedades intencionales devendrá más informativa y más adecuada a la hora de explicar el origen causal de movimientos corporales que la mención de los realizadores físicos que de hecho son responsables causales de tales movimientos. La conclusión, así pues, es que los dos argumentos escépticos nos fuerzan sólo a concluir que las propiedades intencionales son causalmente impotentes con respecto a la conducta pero no a que son causalmente irrelevantes o no pertinentes, algo que, según Jackson y Pettit, reconcilia los argumentos con nuestras intuiciones acerca de la pertinencia causal de lo psicológico con respecto a la conducta.[9]

9. No pondré en cuestión esta opinión aquí aunque, como a la mayoría de filósofos de la mente, me parece muy discutible.

3. Algunas objeciones y un nuevo análisis de la noción de «explicación programática»

A continuación voy a reformular la propuesta de Jackson y Pettit de un modo algo más formal con la intención de que el lector siga con más facilidad tanto mis objeciones a la misma como el análisis alternativo que voy a proponer de la noción de explicación programática.

Digamos, siguiendo a Jackson y Pettit, que una propiedad P «programa» un efecto E cuando y sólo cuando todo realizador de P (o toda propiedad que pertenece a su base de superveniencia) es causalmente eficaz para P. Cuando esto ocurre, la ejemplificación de P asegura la ejemplificación de una propiedad causalmente responsable de E (pues la ejemplificación de la propiedad sobreviniente requiere siempre la ejemplificación de una propiedad de su base de superveniencia). Éste es el sentido metafórico en el que Jackson y Pettit hablan de una propiedad como «programadora» de un efecto.

Llamemos ahora «explicación programática de un efecto E» a cualquier explicación de E en la que se menciona de modo ineliminable una propiedad que programa E, y llamemos «propiedad causalmente pertinente para un efecto E» a aquella cuya mención nos da información causal acerca de E. La tesis central de Jackson y Pettit puede resumirse entonces como la tesis de que una propiedad es causalmente pertinente para E cuando o bien es causalmente eficaz para E o bien programa E.

Paso a continuación a mis objeciones. Mi primera objeción no es una objeción a la propuesta de Jackson y Pettit, sino a la tesis de que tal propuesta constituye un buen alivio para las consecuencias que supuestamente emanan del argumento del duplicado acerca de la eficacia causal de las propiedades psicológicas externistas. Jackson y Pettit afirman que si bien ese argumento muestra que las propiedades psicológicas externistas no son causalmente eficaces para la conducta con todo no debe concluirse que no son causalmente pertinentes, pues las explicaciones de la conducta en términos de tales propiedades son en realidad explicaciones programáticas de la conducta.[10] Ahora bien, esto simplemente no es así. Según el análisis de Jackson y Pettit no es el caso que una propiedad psicológica externista programe una conducta (digamos un determinado movimiento corporal MC). Una propiedad programa un efecto E si y sólo

10. Jackson y Pettit (1988), pp. 398-399.

si todos los realizadores de esa propiedad (o todas las propiedades que pertenecen a la base de superveniencia de esa propiedad) son causalmente eficaces para E. Ahora bien, las propiedades sobre las que sobreviene una propiedad psicológica externista, a saber un complejo de propiedades neurofisiológicas del agente junto con ciertas propiedades del entorno del agente, no son causalmente eficaces para el movimiento MC, sólo el complejo de propiedades neurofisiológicas lo son (de acuerdo con lo que se asume en el argumento del duplicado y Jackson y Pettit aceptan). Por consiguiente, no es cierto que todo realizador de una propiedad psicológica externista cause un movimiento corporal como MC —de hecho, ninguno lo hace—, con lo que la propiedad psicológica externista no programa MC. Al no ser tampoco causalmente eficaz para MC, la conclusión que obtenemos (aplicando el análisis de Jackson y Pettit) es que una propiedad psicológica externista no es causalmente pertinente para un movimiento corporal. De este modo vemos cómo la noción de explicación programática desarrollada por Jackson y Pettit no nos permite aceptar el argumento del duplicado y a la vez defender que las propiedades psicológicas externistas son causalmente pertinentes con respecto a los movimientos corporales.

Mi segunda objeción se refiere a la propuesta misma. El modo en que Jackson y Pettit tratan de reconciliar los argumentos escépticos con nuestra creencia usual de que el contenido intencional es causalmente pertinente para la conducta hace uso de la noción de explicación programática y se articula en dos pasos. En primer lugar, argumentan que toda explicación programática es una explicación causal y, en segundo lugar, argumentan que las explicaciones de la conducta en términos de propiedades funcionales intencionales son explicaciones programáticas. Lo que yo defenderé en cambio es, en primer lugar, que la noción de explicación programática desarrollada por Jackson y Pettit no es apropiada para sustentar la primera premisa y, en segundo lugar, que una vez se desarrolla una noción apropiada para ello, entonces es la segunda premisa la que no puede mantenerse.

Volvamos al ejemplo del ascensor. Estoy totalmente de acuerdo con Jackson y Pettit en que la explicación del hecho de que el ascensor se pare proporcionada a través de la mención de la propiedad del ascensor de tener más de diez personas a bordo es, al menos en un cierto sentido, mejor y más apropiada que la explicación que apela a la propiedad del ascensor de tener veinte personas a bordo. También estoy de acuerdo en que tal cosa puede sostenerse de modo consistente incluso si, movidos por consideracio-

nes similares a las que se exponen en los dos argumentos escépticos, aceptamos que la propiedad mencionada en la primera de las explicaciones es causalmente impotente para el paro del ascensor. También comparto la idea de que tiene perfecto sentido hablar de la propiedad del ascensor de tener más de diez personas a bordo como causalmente pertinente para el efecto puesto que la mención de esta propiedad deviene crucial en ciertas explicaciones causales de ese efecto. Donde discrepo es en el diagnóstico que estos autores hacen de los casos donde hay pertinencia causal sin haber eficacia causal. A la pregunta de por qué la mención de la propiedad del ascensor de tener más de diez personas a bordo proporciona información causal acerca del paro del ascensor, la respuesta de Jackson y Pettit es, como hemos visto, que esta propiedad «programa» el paro del ascensor, esto es, que todo realizador de esa propiedad (la propiedad de tener once personas a bordo, etc.) es causalmente eficaz para el paro del ascensor. A mi juicio, esta respuesta es incorrecta. Es incorrecta porque es incompleta. La respuesta silencia otro hecho crucial que el análisis de Jackson y Pettit ha pasado por alto.

Para mostrar la incorrección del análisis de Jackson y Pettit, partiré del supuesto de que una explicación causal de un efecto E es una explicación en la que se proporciona información acerca del origen causal de E.[11] Ahora bien, cuando explicamos el paro del ascensor apelando a la propiedad suya de tener más de diez personas a bordo sabemos que lo que decimos puede ser verdadero sólo si es verdad que el ascensor se para si suben a él once personas, o doce, o trece, etc. De este modo, la mención de la propiedad programadora nos proporciona información acerca de las propiedades que de veras son causalmente eficaces para el paro del ascensor, y en definitiva nos proporciona información causal sobre el paro del ascensor; si se quiere, nos proporciona información acerca de orígenes causales posibles del paro del ascensor. Lo que es resaltable de esta reflexión es que sabemos que el enunciado «el ascensor se paró porque subieron a él más de diez personas» es verdadero sólo si es verdadero que el ascensor se para si suben a él once personas, o doce, o trece, etc., como consecuencia de nuestro conocimiento acerca de las condiciones de verdad del enunciado citado. Es decir, este co-

11. Véase Lewis (1986). En realidad, Lewis sostiene algo más fuerte, a saber, que cualquier explicación de la ejemplificación de una propiedad debe dar información causal sobre el origen de su *explanandum*. No es necesario, no obstante, compartir esta tesis más fuerte de Lewis, basta para mis propósitos con el supuesto que aparece en el texto principal.

nocimiento es parte o, si se prefiere, una consecuencia de nuestra comprensión del enunciado, esto es, del hecho de que entendemos su significado.

¿Por qué es esto así? La respuesta es que es así porque tener once personas a bordo, doce personas a bordo, etc., son todos ellos determinados del determinable tener más de diez personas a bordo, y esto significa que se dan ciertas relaciones de implicación conceptual entre esas propiedades. En primer lugar, la ejemplificación de cualquier determinado, como tener veinte personas a bordo, implica conceptualmente la ejemplificación del determinable tener más de diez personas a bordo, es decir, incurriría en un error conceptual quien dijera que un objeto ejemplifica el determinado y en cambio no ejemplifica el determinable. En segundo lugar, la ejemplificación del determinable tener más de diez personas por parte de un objeto implica conceptualmente la ejemplificación de uno de sus determinados por parte de ese mismo objeto. Dicho de otro modo: es conceptualmente imposible que algo X ejemplifique la propiedad de tener veinte personas a bordo y no ejemplifique la propiedad de tener más de diez personas a bordo y también es conceptualmente imposible que algo X ejemplifique la propiedad de tener más de diez personas a bordo y no ejemplifique la propiedad de tener once personas a bordo, ni tampoco la propiedad de tener doce personas a bordo, etc.

Llamemos a tales implicaciones «implicaciones P». La razón por la que el enunciado «el ascensor se paró porque más de diez personas subieron a bordo» transmite la información de que el ascensor se para si suben once personas, o doce personas, etc., es que la propiedad de tener más de diez personas a bordo mantiene las relaciones de implicación-P con las propiedades de subir a bordo once personas, doce personas, etc. La existencia de estas implicaciones-P es lo que explica que la mención del determinable —tener más de diez personas a bordo— proporcione información acerca de las propiedades causalmente eficaces para el efecto, a saber, sus determinados, dándonos así información causal valiosa acerca de ese efecto.

Lo que es importante en esta explicación de por qué propiedades causalmente impotentes para un efecto pueden ser no obstante causalmente pertinentes para ese mismo efecto, es que en ausencia de estas relaciones de implicación conceptual que hemos llamado «implicaciones-P» entre la propiedad sobreviniente y las propiedades subvinientes (aquellas que pertenecen a su base de superveniencia), la mención de la propiedad sobreviniente no va a propor-

cionar información causal sobre el efecto porque no va a darnos información acerca de las propiedades subvinientes, y en ese caso no obtendremos a través de dicha mención ningún tipo de información causal contrafáctica sobre el efecto. Ahora bien, si tal mención no proporciona información causal acerca del efecto no debería por tanto considerarse causalmente pertinente para el efecto.

A la luz de lo que acabamos de observar voy a proponer a continuación un análisis de la noción de explicación programática que tenga lo dicho en cuenta. Diremos que una propiedad Q programa un efecto E si y sólo si (1) todo realizador de Q es causalmente eficaz para E, y (2) Q mantiene con sus realizadores las relaciones de implicación-P.

Lo que he argumentado anteriormente es simplemente que sólo cuando se añade la cláusula (2) a la caracterización de una propiedad programadora al análisis original de Jackson y Pettit —que sólo contempla la cláusula (1)— puede mantenerse que las explicaciones programáticas de un efecto, esto es, explicaciones en las que se menciona de modo ineliminable una propiedad programadora del efecto, son todas ellas explicaciones causales de ese efecto. Esto es, la mención de una propiedad sobreviniente Q proporcionará información acerca de sus propiedades subvinientes causalmente eficaces, sólo cuando entre las propiedades causalmente eficaces y la propiedad sobreviniente Q existan las relaciones de implicación-P. En ausencia de tales relaciones de implicación-P, así pues, la mención de la propiedad sobreviniente Q no nos dirá nada acerca de las propiedades subvinientes causalmente eficaces, y por consiguiente no proporcionará información causal alguna.[12]

Esto es exactamente lo que ocurre con las explicaciones funcionales de efectos que aparecen especificados en el rol funcional asociado a la propiedad funcional que aparece en el *explanans*. El problema en estos casos es que las propiedades funcionales no mantienen con sus realizadores las relaciones de implicación conceptual-P. Consideremos otra vez el caso de la solubilidad en agua a modo de ejemplo. Que algo X ejemplifique la solubilidad en agua sólo implica conceptualmente que X ejemplifica *alguna* propiedad que causa la disolución en caso de inmersión en agua, pero no implica ningún realizador suyo.[13] Supongamos que ser DIS es una de las propiedades microestructurales causalmente responsables

12. Ello siempre bajo el supuesto, compatible con los argumentos escépticos, de que la propiedad sobreviniente misma no es causalmente eficaz para el efecto a explicar.

13. Al menos no los implica bajo la descripción pertinente de los mismos. Sobre este punto véase la sección siguiente.

de la disolución en caso de inmersión en agua (es decir, supongamos que ser DIS es uno de los realizadores de la solubilidad en agua) y supongamos también que algo X es soluble en agua. Pues bien este último hecho no implica conceptualmente nada acerca de la ejemplificación de ser DIS. Por otro lado, supongamos que algo X ejemplifica ser DIS. Nuevamente este hecho no implica conceptualmente nada acerca de que X sea soluble en agua. La implicación de «X es DIS» a «X es soluble en agua» precisa de una premisa *a posteriori* y no conceptual acerca de los poderes causales de ser DIS (en concreto, entre los poderes causales de ser DIS debe encontrarse el de causar la disolución del objeto en que se ejemplifica si ese objeto es sumergido en agua). Ésta es la razón por la que afirmar que X se disolvió en agua porque era soluble en agua no nos dice nada acerca de los procesos causales contrafácticos que llevan a la disolución de X en agua y, por consiguiente, no debería considerarse tal explicación una explicación causal de su *explanandum*.

Es un lugar común que la explicación de la disolución en agua de un objeto determinado X que menciona de modo indispensable el hecho de que X ejemplifica la propiedad de ser soluble en agua no es una explicación muy iluminadora. En realidad yo me comprometo con algo más fuerte, a saber, que una explicación tal no puede considerarse una explicación causal en absoluto (si uno está dispuesto, como lo están Jackson y Pettit, a no considerar a la solubilidad en agua una propiedad causalmente eficaz para la disolución). Insisto en que en un caso así, y contra lo que quieren sostener Jackson y Pettit, no se da ningún tipo de información causal contrafáctica acerca del efecto. La única información transmitida por el enunciado 'X se disolvió en agua porque era soluble en agua' (excluyendo la eficacia causal de la solubilidad en agua) es que alguna propiedad ejemplificada por X causó la disolución de X. Nos dice solamente que algún tipo de proceso causal ha tenido lugar pero no nos da ninguna pista acerca de qué proceso causal ha tenido lugar, o acerca de qué procesos causales podrían haber tenido lugar. Sólo nos dice que es pertinente ofrecer una explicación causal de la disolución de X en términos de alguna propiedad ejemplificada por X pero, hablando con propiedad, no nos da una explicación causal de este tipo.

Jackson y Pettit discrepan de este punto de vista. Ciertamente reconocen que ciertas explicaciones funcionales, como la de nuestro ejemplo acerca de la solubilidad en agua, son menos informativas que otras explicaciones programáticas, como la de nuestro

ejemplo del ascensor.[14] Esta observación resulta sorprendente por cuanto según el análisis ofrecido por Jackson y Pettit la noción de explicación programática no admite grados. Recuérdese que según ese análisis una explicación programática de un efecto E es una explicación en la que se menciona en el *explanans* una propiedad tal que todas las propiedades de su base de superveniencia son causalmente eficaces para E. Pero esto se cumple igual de bien tanto en el caso de la solubilidad en agua con respecto a la disolución como en el de la propiedad de tener más de diez personas a bordo con respecto al paro del ascensor. Esto sugiere una vez más que el análisis que ofrecen Jackson y Pettit es incompleto. Mi tesis es que, una vez que se lo completa de modo satisfactorio, las explicaciones funcionales del tipo de la del ejemplo de la solubilidad no pueden ser consideradas explicaciones programáticas en absoluto.[15]

Si mi objeción al análisis de la noción de explicación programática ofrecido por Jackson y Pettit es correcta, entonces se derivan de ella importantes consecuencias con respecto a la posición que estos autores defienden frente a los dos argumentos escépticos mencionados al comienzo de esta sección. Recuérdese que Jackson y Pettit sostienen que si bien esos argumentos muestran que las propiedades intencionales no son causalmente eficaces para los movimientos corporales, no muestran que no sean causalmente pertinentes. Las propiedades intencionales serían de hecho causalmente pertinentes porque programan ciertos movimientos corporales (aquellos implicados en las acciones intencionales). Sin embargo, bajo el nuevo análisis de la noción de explicación programática que acabo de proponer esto ya no puede mantenerse. Las propiedades funcionales, al menos si se las considera como propiedades funcionales tales que los roles funcionales asociados a ellas incluyen una especificación de movimientos corporales, no son propiedades que programen esos movimientos corporales por la sencilla razón de que no satisfacen la segunda cláusula del nuevo análisis de la noción de explicación programática. Es decir, una propiedad intencional no mantiene con sus supuestos realizadores físicos las relaciones de implicación conceptual-P (¡si así fuera sería una verdad conceptual que la ejemplificación de ciertas propiedades psicológi-

14. Véase Jackson y Pettit (1988), p. 396. Lo que dicen exactamente es que ciertas explicaciones funcionales son más informativas que otras. En la siguiente sección mostraré cómo el análisis que propongo de la noción de explicación programática da perfecta cuenta de esta intuición de Jackson y Pettit.

15. Aunque otro tipo de explicaciones funcionales sí van a ser consideradas como tales. Véase, de nuevo, la sección siguiente.

cas implica la ejemplificación de ciertas otras propiedades físicas!). Ésta es de nuevo la razón por la cual afirmar que un determinado movimiento corporal MC tendente a la consecución de un estado de cosas A fue causado por, digamos, un deseo de A no nos dice nada acerca de procesos causales físicos contrafácticos que levan a MC, y por consiguiente no debe ser considerada tal afirmación una explicación causal de MC.[16]

Lo que esta objeción muestra es que la posición de Jackson y Pettit con respecto al contenido funcional restringido es errónea. El contenido funcional restringido no debe ser considerado causalmente pertinente con respecto a la conducta, no al menos si por «causalmente pertinente» se entiende indispensable en ciertas explicaciones causales de la conducta. Como vimos al comienzo de esta sección, exactamente lo mismo vale para el contenido funcional externista. De este modo llegamos a la principal conclusión negativa del presente trabajo, a saber, que el descubrimiento valioso de las explicaciones programáticas por parte de Jackson y Pettit es lamentablemente insuficiente para defender la pertinencia causal del contenido psicológico con respecto a la conducta frente a los dos argumentos escépticos mencionados al comienzo.

4. Variedades de explicaciones funcionales y algunos refinamientos del nuevo análisis

El propósito de esta sección es refinar el análisis de la noción de explicación programática propuesto en la sección anterior. Algunos de los refinamientos que voy a proponer vienen motivados por la consideración de algunos casos de explicaciones funcionales y de explicaciones que apelan a determinables que no son cubiertos por el análisis propuesto hasta ahora. Otros obedecen a ciertos detalles técnicos que he omitido hasta ahora para facilitar la comprensión de mi crítica a Jackson y Pettit pero que es necesario tener en cuenta para aumentar el grado de precisión del análisis. En cualquier caso, las ideas centrales del mismo discutidas en la sección anterior se mantendrán intactas.

Hasta ahora hemos tomado en consideración solamente casos en los que se apela a una propiedad funcional para explicar la ejem-

16. Una vez más, téngase en cuenta que tal conclusión sólo se sigue si decidimos aceptar los argumentos escépticos y consideramos que las propiedades psicológicas intencionales no son causalmente eficaces para los movimientos corporales.

plificación de una propiedad categórica (no funcional) especificada en el rol funcional asociado a ella. Es éste el caso de la solubilidad en agua y la disolución en agua o de los deseos y los movimientos corporales. Vamos ahora a enfretarnos a dos casos que no responden a esta estructura.

Según la mayoría de análisis funcionalistas, una propiedad funcional puede aparecer mencionada en la especificación del rol funcional asociado a otra propiedad funcional. De acuerdo con muchos filósofos de la mente funcionalistas esto es justamente lo que ocurre con las propiedades intencionales. Así, se suele decir que creer que P y Q son el caso es una propiedad parcialmente caracterizada por el hecho de que causa en quien la ejemplifique la tendencia a inferir a causa de ella que P es el caso. Dado que este modo usual de formular el punto de vista funcionalista da precisamente por sentado aquello que es objeto de discusión en este trabajo, a saber, la pertinencia causal de las propiedades funcionales, sugiero reformular tal punto de vista como la tesis de que ejemplificar la propiedad de creer que P y Q son el caso requiere ejemplificar otra propiedad que causa la ejemplificación de un realizador de la propiedad de creer que P es el caso. La cuestión ahora es si puede sostenerse que creer que P y Q es el caso es una propiedad programadora de creer que P es el caso.

De acuerdo con el análisis alternativo al de Jackson y Pettit sugerido al final de la sección anterior la respuesta es rotundamente negativa. La mención de la propiedad de creer que P y Q son el caso no nos informa acerca de cuáles son sus realizadores (físicos), a lo sumo nos dice que tiene realizadores, pero no nos da ninguna pista acerca de cuáles son. La razón de ello una vez más es que se incumple también en este caso la cláusula (2) del nuevo análisis.

Ahora bien, sí parece haber casos genuinos de explicación causal que involucran determinables y que responden a esta misma estructura. Esto es, en algunos casos parece darse información causal acerca de un determinable en términos de otro determinable. Un caso de este tipo viene proporcionado por los propios Jackson y Pettit.[17] Considérense dos electrones A y B sobre los que actúan dos fuerzas independientes F_A y F_B respectivamente. y supóngase que el electrón A acelera según la misma razón que B. Una explicación perfectamente plausible de este hecho es que la magnitud de ambas fuerzas es la misma.

Un posible análisis de la situación del ejemplo es el siguiente. La ejemplificación por parte de A y de B de la propiedad relacional

17. Jackson y Pettit (1988), pp. 392-393.

de ser sometidos a fuerzas de la misma magnitud es invocada para explicar la ejemplificación por parte de A y de B de la propiedad relacional de acelerar según la misma razón. Un realizador de la primera propiedad, uno de sus determinados, es la ejemplificación por parte tanto de A como de B de la propiedad (no relacional) de ser sometido a una fuerza de magnitud *n*; y un realizador de la segunda propiedad relacional es también uno de sus determinados, digamos la ejemplificación tanto por parte de A como de B de la propiedad de acelerar según la razón *m*. Las conexiones causales se dan aquí entre la propiedad no relacional de ser sometido a una fuerza de magnitud *n* y la propiedad no relacional de acelerar según la razón *m*, pero no es necesario suponer una conexión causal entre los realizadores (o determinados) de las propiedades relacionales ni tampoco entre el determinable cuya ejemplificación se quiere explicar ni el determinable al que se apela para explicarlo.

Bajo este análisis, que es el que considero más plausible, la propiedad relacional que aparece en el *explanans* no programa la propiedad relacional que aparece en el *explanandum*, ni de acuerdo con mi definición ni tampoco de acuerdo con la definición de Jackson y Pettit, en contra de lo que ellos parecen suponer, porque para que una propiedad programe un efecto se requiere en ambos casos que todo realizador de esa propiedad cause el efecto.

Si aceptamos los argumentos escépticos discutidos anteriormente y también el análisis que he propuesto para este ejemplo, según el cual no existe conexión causal entre los determinados de las dos propiedades relacionales, entonces no parece ser posible una explicación causal de la ejemplificación de la propiedad relacional de acelerar según la misma razón por parte de A y de B por la sencilla razón de que la ejemplificación de esta propiedad no constituye un efecto causal. Aun así, estimo conveniente modificar el análisis de modo que dé cuenta de casos como éste. La razón para ello es la siguiente. Por un lado, la relación entre las dos propiedades relacionales, si bien no parece causal, sí sobreviene sobre relaciones causales entre propiedades no relacionales que son constituyentes de sus realizadores. Por otro lado, las propiedades relacionales mantienen con sus realizadores las relaciones de implicación-P. Todo ello hace posible que la mención del proceso sobreviniente nos dé una información importante acerca de procesos causales que involucran constituyentes de las propiedades subvinientes. Propongo llamar a este tipo de explicación, a falta de un nombre mejor, explicación «paracausal» y entiendo que, debido a las razones que acabo de exponer, es plausible considerarla como un tipo de explicación programática.

Pero para poder considerarla como tal debemos introducir una modificación en el análisis anterior, a saber: una propiedad Q programa un efecto E si y sólo si (1) todo realizador de Q es causalmente eficaz para E, y (2) Q mantiene con sus realizadores las relaciones de implicación-P; o bien, (1) para cada realizador de Q las propiedades que lo constituyen son causalmente eficaces para las propiedades que constituyen algún realizador de E, y (2) Q y E mantienen con sus realizadores la relaciones de implicación-P.

El siguiente caso que paso a considerar es aquel en que se apela a una propiedad funcional para explicar un efecto que no aparece especificado en el rol funcional asociado a ella. Frank Jackson ha dado un ejemplo que responde a esta estructura.[18] Así, Richard Feynman explicó la explosión del Challenger en términos de la pérdida de elasticidad a bajas temperaturas de uno de los componentes de la nave. Para Jackson esto tiene la apariencia de una explicación perfectamente plausible y yo estoy totalmente de acuerdo con él. Estoy de acuerdo porque esta explicación nos proporciona información acerca del proceso causal que de hecho llevó a la explosión del Challenger. En este caso el *explanans* contiene de modo indispensable una mención a una propiedad funcional, la pérdida de elasticidad. El reconocimiento de estos casos, no obstante, no refuta mi crítica al análisis de las explicaciones programáticas propuesto por Jackson y Pettit, sino que más bien la confirma. Nótese que la pérdida de elasticidad puede caracterizarse como esa propiedad funcional que requiere para su ejemplificación en un objeto de la ejemplificación de una propiedad que cause la falta de expansión de ese objeto. Esta última propiedad, así pues, mantiene con la pérdida de elasticidad el tipo de implicaciones-P exigidos en mi análisis. Es por ello que al citar la pérdida de elasticidad de un componente de la nave como una explicación de la explosión la explicación sólo puede entenderse como implicando la afirmación de que el hecho de que este componente perdiera su capacidad para expandirse está en el origen causal del desastre del Challenger. En consecuencia en este caso la mención de una propiedad funcional nos dice algo acerca del proceso causal que de hecho llevó a la explosión por la misma razón que lo hace la mención de cualquier propiedad programadora, a saber, porque mantiene las relaciones de implicación-P con una propiedad causalmente eficaz para esa explosión.

18. Jackson (1996), pp. 396-397.

De nuevo, tener en cuenta estos casos como casos de explicación programática requiere una nueva modificación en el análisis. He aquí esta nueva modificación:

(PP) Una propiedad Q programa un efecto E si y sólo si (1) todo realizador de Q es causalmente eficaz para E, y (2) Q mantiene con sus realizadores o con alguna o algunas propiedades causalmente eficaces para E las relaciones de implicación-P; o bien, (1) para cada realizador de Q las propiedades que lo constituyen son causalmente eficaces para las propiedades que constituyen algún realizador de E, y (2) Q y E mantienen con sus realizadores la relaciones de implicación-P.

Siguiendo en este punto a Jackson y Pettit podemos entonces caracterizar la noción de propiedad causalmente pertinente como sigue:

(PCP) Una propiedad P es causalmente pertinente para un efecto E si y sólo si (1) P es causalmente eficaz para E; o (2) P programa E.

Acepto, como sugieren Jackson y Pettit, que la mención de algunas propiedades puede proporcionar información causal acerca de un efecto aunque esas propiedades mencionadas no sean causalmente responsables del mismo, pero en mi opinión, como he tratado de argumentar hasta aquí, eso ocurre sólo cuando esas propiedades mantienen con las que realmente son causalmente responsables del efecto las relaciones conceptuales de implicación-P.

Antes de pasar a las conclusiones es necesario hacer dos observaciones. La primera de ellas es técnica. Debe tenerse en cuenta algo que por razones de claridad expositiva no he mencionado hasta ahora, a saber, que una propiedad mantiene con otra las relaciones conceptuales de implicación-P sólo según como se las describa a ambas. Por ejemplo algunos funcionalistas defienden que las propiedades funcionales son simplemente una disyunción de sus realizadores.[19] Supongamos, por mor de la discusión, que esta teoría metafísica acerca de las propiedades funcionales es correcta. En ese caso resulta que cada propiedad funcional, si se la describe como una disyunción de sus realizadores, mantendrá con sus realizadores las re-

19. Sobre propiedades disyuntivas, véase Owens (1989), Fodor (1998) y Clapp (2001).

laciones de implicación-P.[20] Así pues resulta que, de acuerdo con nuestro análisis de propiedad programadora, las propiedades funcionales programan los efectos de sus realizadores después de todo.

Frente a este razonamiento deberíamos tener bien presente que las explicaciones generan contextos intensionales. En consecuencia, la mención de una propiedad resultará causalmente informativa para un efecto E en el contexto de una explicación de E sólo si la mención de esa propiedad nos proporciona información acerca del origen causal de E *tal y como se la describe en esa explicación*. Lo que hemos mostrado es que una propiedad funcional no proporciona este tipo de información acerca de un efecto especificado en el rol funcional asociado a ella *cuando se la describe funcionalmente*, esto es, como aquella propiedad que se ejemplifica al ejemplificarse una propiedad que tenga el rol funcional asociado. Similarmente, podría sostenerse que una propiedad funcional F, descrita funcionalmente, implica conceptualmente la disyunción de sus realizadores, descritos como «realizadores de F», pero adviértase que la mención de los realizadores de F bajo esa descripción no nos da ninguna información acerca de los procesos causales que involucran a esos realizadores.

No deja de ser cierto, no obstante, que siendo estrictos deberíamos una vez más modificar la definición (PP) de modo que la definición modificada tenga en cuenta que la relación entre propiedades consistente en mantener las implicaciones-P es sensible al modo como esas propiedades se describan. Es importante señalar que la dificultad técnica que eso entraña es general y no concierne únicamente a las propiedades funcionales. Volvamos para ver esto al ejemplo original del ascensor y supongamos que la propiedad de tener a bordo más de diez personas es la propiedad favorita de Jackson y Pettit. Considérese ahora el siguiente enunciado: «el ascensor se paró porque ejemplificó la propiedad favorita de Jackson y Pettit». Obsérvese que este enunciado no transmite información causal contrafáctica alguna acerca del paro del ascensor a pesar de que la propiedad mencionada en el *explanans* es precisamente la misma que la mencionada en «el ascensor se paró porque subieron a él más de diez personas».

La última modificación del análisis de propiedad programadora que voy a proponer en este trabajo trata de tener en cuenta este hecho del modo más simple posible:

20. Por supuesto, esto será así cuando a su vez a los realizadores se los describa con los mismos predicados que aparecen en la disyunción.

(PP') Una propiedad Q, bajo una descripción *d*, programa un efecto E, descrito bajo una descripción *d'*, si y sólo si (1) todo realizador de Q es causalmente eficaz para E, y (2) Q, bajo la descripción *d*, mantiene con sus realizadores o alguna propiedad causalmente eficaz para E, bajo algunas descripciones de los mismos, las relaciones de implicación-P; o bien, (1) para cada realizador de Q las propiedades que lo constituyen son causalmente eficaces para las propiedades que constituyen algunos de los realizadores de E, y (2) Q, bajo la descripción *d*, y E, bajo la descripción *d'*, mantienen con sus realizadores, bajo alguna descripción de los mismos, las relaciones de implicación-P.

Sobre esta caracterización conviene tener presentes dos cosas. Primero, una explicación programática será considerada correcta o no en la medida en que las propiedades mencionadas en las descripciones de los realizadores bajo las cuales la propiedad programadora, tal como aparece descrita en el *explanans*, mantiene las relaciones de implicación-P sean o no causalmente eficaces para el *explanandum*. Segundo, de acuerdo con el análisis modificado (PP') diremos que una propiedad P, bajo una descripción *d*, es causalmente pertinente para un efecto E, bajo una descripción *d'*, si y sólo si P es causalmente eficaz para E, o bien P, bajo *d*, programa E bajo *d'*.

La segunda observación que quería hacer viene motivada por una reflexión sobre ejemplos como el del Challenger. Estos ejemplos muestran que algunas explicaciones funcionales son explicaciones programáticas, esto es, que la mención de una propiedad funcional (descrita funcionalmente) puede proporcionar información acerca del origen causal del efecto sin ser causalmente responsable del mismo. Sin embargo, hemos visto también que muchas supuestas explicaciones funcionales de un efecto no son explicaciones programáticas pues en ellas la mención de la propiedad funcional no nos dice nada acerca del origen causal del efecto. Esto creo que nos da pie a trazar la siguiente distinción entre dos tipos de explicación funcional:

(EF1) Una explicación funcional de un efecto E en términos de una propiedad funcional F, descrita funcionalmente, es una explicación programática de E si y sólo si alguna de las propiedades especificadas en el rol funcional asociado a F es causalmente responsable de E.

(EF2) Una explicación funcional de un efecto E en términos de una propiedad funcional F, descrita funcionalmente, no es una ex-

plicación programática de E si E involucra la ejemplificación de una propiedad especificada en el rol funcional asociado a F.

Como corolario, nótese que se sigue algo importante de (EF1) y (EF2) y de los dos argumentos escépticos discutidos en las secciones precedentes. Lo que se sigue es que una propiedad intencional, tanto si se la analiza de modo externista como de modo restringido, nunca es causalmente pertinente, bajo una descripción funcional de la misma, para un movimiento corporal a menos que se considere plausible un análisis funcionalista de las propiedades intencionales según el cual no debe hacerse ninguna referencia a movimientos corporales, ni a nada que involucre movimientos corporales, en la especificación de los roles funcionales asociados a ellas.

Una última observación. La distinción entre explicaciones funcionales que cuentan como explicaciones programáticas y las que no cuentan como tales nos ayuda en mi opinión a clarificar por qué se perciben ciertas explicaciones funcionales como más informativas que otras. Todos estamos de acuerdo en que decir que X se disolvió porque era soluble en agua es poco informativo (si es que realmente transmite alguna información); ahora bien, al decir que X explotó a causa de la pérdida de elasticidad de un componente de X sí parecemos transmitir información causal valiosa acerca del efecto a explicar. La elucidación de la noción de explicación programática que he llevado a cabo en este trabajo explica estas intuiciones. Nos permite decir que una explicación funcional del segundo tipo es informativa precisamente porque es una explicación programática, mientras que la segunda no lo es y ello explica su escaso o nulo carácter informativo; nos permite en suma comprender por qué algunas explicaciones funcionales son informativas mientras que otras no lo son. Esto es sólo una muestra de la fertilidad filosófica que puede tener el análisis refinado de la noción de explicación programática que acabo de proponer.

Referencias bibliográficas

ARMSTRONG, D., *A Materialist Theory of the Mind*, Routledge, Londres, 1968 (edición revisada, 1993).

BLOCK, N., «Troubles with Functionalism», en SAVAGE, C. W. (ed.), *Perception and Cognition. Issues in the Foundations of Psychology*, University of Minnesota Press, Minneapolis, 1978, pp. 261-325.

BURGE, T., «Individualism and the Mental», *Midwest Studies in Philosophy,* v. 4, (1979), pp. 73-121.

CLAPP, L., «Disjunctive Properties: Multiple Realizations», *Journal of Philosophy*, v. 98, (2001), pp. 111-136.

FODOR, J., «Special Sciences: Still Autonomous After All These Years», en FODOR, J., *In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind*, MIT Press, Cambridge, 1998, pp. 1-12.

HORGAN, T., «From Supervenience to Superdupervenience: Meeting the Demands of a Material World», *Mind*, v. 102, (1993), pp. 555-586.

JACKSON, F., «Mental Causation», *Mind*, v. 105 (1996), pp. 377-413.

JACKSON, F. y PETTIT, P., «Functionalism and Broad Content», *Mind*, v. 97, (1988), pp. 381-400.

JACKSON, F. y PETTIT, P., «Causation in the Philosophy of Mind», *Philosophy and Phenomenological Research*, v. 50, suplemento (1990), pp. 195-214.

KIM, J., «Concepts of Supervenience», *Philosophy and Phenomenological Research*, v. 45, (1984), pp. 153-176.

KIM, J., «The Nonreductivist's Troubles with Mental Causation», en HEIL, J. y MELE, A. (eds.), *Mental Causation*, Clarendon Press, Oxford, 1993, pp. 189-210.

KIM, J., *Mind in a Physical World*, MIT Press, Cambridge (MA), 1998.

LEWIS, D., «An Argument for the Identity Theory», *Journal of Philosophy*, v. 63, (1966), pp. 17-25.

LEWIS, D., «Events», en LEWIS, D., *Philosophical Papers. Volume II*, Oxford University Press, Nueva York, 1986, pp. 241-69.

OWENS, D., «Disjunctive Laws», *Analysis*, v. 49, (1989), pp. 197-202.

PUTNAM, H., «The Mental Life of Some Machines», en PUTNAM, H., *Mind, Language and Reality*, Cambridge University Press, Cambridge, 1967/1975, pp. 408-428.

PUTNAM, H., «Philosophy and our Mental Life», en PUTNAM, H., *Mind, Language and Reality*, Cambridge University Press, Cambridge, 1973/1975, pp. 291-303.

PUTNAM, H., «The Meaning of 'Meaning'», en PUTNAM, H., *Mind, Language and Reality*, Cambridge University Press, Cambridge, 1975, pp. 215-271.

RYLE, G., *The Concept of Mind*, Barnes and Noble, N. York, 1949.

SCHIFFER, S., *Remnants of Meaning*, MIT Press, Cambridge (MA), 1987.

ÍNDICE TEMÁTICO

ÍNDICE DE NOMBRES

COLABORADORES

José A. Díez Calzada. Profesor Titular de Lógica y Filosofía de la Ciencia en la Universidad Rovira y Virgili de Tarragona. Realizó la Tesis Doctoral bajo la dirección de C. Ulises Moulines. Con este autor ha publicado diversos trabajos, como «Theories as Nets: Combinatorial Measurement Theory» (1994) y *Fundamentos de Filosofía de la Ciencia* (1997). La Concepción estructural ha sido su línea de investigación prefente. Entre sus publicaciones recientes figuran «Hacia una Teoría General de la Representación Científica» (1998) y «Structuralist Analysis of Fundamental Measurement Theories» (2000).

Javier Echeverría. Catedrático de Ciencia, Tecnología y Sociedad del CSIC, Premio Nacional de Ensayo año 2000 y Premio Euskadi de Investigación. Ha sido Presidente de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España. Ha publicado importantes trabajos tanto en el ámbito de la Historia de la Ciencia como en la esfera de la Filosofía de la Ciencia y de la Filosofía de la Tecnología. Ha coordinado el número monográfico *La Filosofía de la Ciencia como Filosofía práctica* (1995). Entre sus libros figuran *Análisis de la identidad* (1987), *Filosofía de la Ciencia* (1995) e *Introducción a la Metodología de la Ciencia* (2 ediciones, 1986 y 1999).

Rafaela García Elskamp. Premio extraordinario de Licenciatura (Universidad de Murcia). Realizó la Tesis Doctoral sobre «La Teoría de la Acción como fundamentación de las Ciencias Humanas y Sociales en la Escuela Finlandesa de Filosofía de la Ciencia». Pertenece al Grupo de Investigación de «Filosofía de la Ciencia-Metodología de la Ciencia» de la Universidad de A Coruña. Entre sus publicaciones están los trabajos «Intención e intencionalidad: Estudio comparativo» (1986), «Intencionalidad y acción social» (1988), «El realismo científico-crítico de R. Tuomela» (1989), «Acción e Historia en la Escuela Finlandesa» (1996).

Wenceslao J. González. Catedrático de Lógica y Filosofía de la Ciencia de la Universidad de A Coruña. Premio de Investigación en Humanidades de Galicia. Ha realizado estancias de investigación en la Universidad de St. Andrews (Escocia), la Universidad de Münster (Alemania), la *London School of Economics* (Inglaterra) y el *Center for Philosophy of Science* de la Universidad de Pittsburgh (Estados Unidos), como *Visiting fellow*. Coordina los volúmenes de *Gallaecia. Estudios de Filosofía y Metodología de la Ciencia*: *Progreso científico e innovación tecnológica* (1997); *El Pensamiento de L. Laudan. Relaciones entre Historia de la Ciencia y Filosofía de la Ciencia* (1998); *Ciencia y valores éticos* (1999); *Problemas filosóficos y metodológicos de la Economía en la Sociedad tecnológica actual* (2000); *La Filosofía de Imre Lakatos: Evaluación de sus propuestas* (2001).

Javier Ordóñez. Catedrático de Lógica y Filosofía de la Ciencia en la Universidad Autónoma de Madrid. Ha editado trabajos de científicos representativos, como las investigaciones de L. Boltzmann *Escritos de Mecánica y Termodinámica* (1986) y el estudio de S. Carnot *Reflexiones sobre la potencia motriz del fuego* (1987). Es también uno de los editores de los libros *La Ciencia y su público* (1990) y *Después de Newton: Ciencia y sociedad durante la primera revolución industrial* (1998). Cuenta con artículos como «Cosmología» (1990), «History of Science in Spain» (1990) y «The Story of a Non-discovery: Helmholtz and the Conservation of Energy» (1996). Es coautor de libros como *Teorías del Universo*, vol. I: *De los pitagóricos a Galileo* (1999), y *Teorías del Universo*, vol. II: *De Galileo a Newton* (1999).

David Pineda Oliva. Profesor de la Universidad de Girona. Doctor en Filosofía por la Universidad de Barcelona con la Tesis «Problemas y perspectivas del fisicismo». Sus investigaciones giran, fundamentalmente, en torno a problemas metafísicos relacionados con la mente, como mente-cuerpo y la causalidad mental. Sobre estos temas ha publicado trabajos en revistas especializadas editadas en nuestro país (*Theoria*, *Teorema*) y en el extranjero (*Crítica*, *Philosophical Issues*, *European Journal of Philosophy*). Prepara actualmente un libro sobre Filosofía de la Mente.

Merrilee H. Salmon. Catedrática de Historia y Filosofía de la Ciencia en la Universidad de Pittsburgh. Ha sido Editora de *Philosophy of Science*, la revista de la *PSA*, la Sociedad americana de Filosofía de la Ciencia. Ha sido Directora del Departamento de *History and Philosophy of Science*. Se ha especializado en la explicación en Ciencias Sociales, habiendo publicado con Wesley Salmon «Alternative Models of Scientific Explanation» (1979). Colabora en el volumen *Introduction to the Philosophy of Science* (1992). Su concepción ha recibido la atención del *Workshop in Honor of Merrilee and Wesley Salmon*, organizado por el Centro de Filosofía e Historia de la Ciencia de Florencia, simposio celebrado en Florencia, Italia, en mayo de 1996.

Wesley C. Salmon. Catedrático de Filosofía de la Universidad de Pittsburgh y ha sido el Presidente de la Sociedad Internacional de Lógica, Metodología y Filosofía de la Ciencia (1996-99). Discípulo directo de Hans Reichenbach, que fue el Director de su Tesis Doctoral, realizada en la Universidad de California en Los Angeles (UCLA). Tras impartir docencia en la UCLA, fue profesor de las Universidades de Washington (State College), Northwestern, Brown, Indiana, Arizona y Pittsburgh. Fue Presidente de la *Philosophy of Science Association* (1971-72); Presidente de la *American Philosophical Association*, Pacific Division (1977-78); y Presidente de la *International Union of History and Philosophy of Science* (1998-99). Escribió libros especialmente influyentes: *Logic* (3 ediciones: 1963, 1973 y 1984, traducido al japonés y al portugués), *The Foundations of Scientific Inference* (1967), *Statistical Explanation and Statistical Relevance* (1971, con contribuciones de Richard C. Jeffrey y James G. Greeno), *Space, Time, and Motion* (2 ediciones: 1975 y 1981), *Scientific Explanation and the Causal Structure of the World* (1984), *Four Decades of Scientific Explanation* (1990, traducido al italiano), y *Causality and Explanation* (1998).

La diversidad de la explicación científica centra el presente volumen: por un lado, analiza qué es explicar en la Ciencia; y, por otro, pone de relieve que la actividad científica está abierta a diversos tipos de explicación. Así, la estructura del libro se articula en varios apartados: "evolución del problema de la explicación científica", "propuestas para las explicaciones científicas", "las explicaciones causales: estructura y aplicación", y "explicaciones teleológico-causales y explicaciones programáticas".

Este libro pertenece a la serie de volúmenes *Gallaecia. Estudios de Filosofía y Metodología de la Ciencia*, que coordina Wenceslao J. González (Catedrático de Lógica y Filosofía de la Ciencia, U. de A Coruña) y que analiza temas actuales en contacto con filósofos de la Ciencia relevantes (L. Laudan, I. Niiniluoto, E. Agazzi, D. Hausman, J. Worrall, ...). Versan sobre cuestiones centrales: "Progreso científico e innovación tecnológica"; "El Pensamiento de L. Laudan. Relaciones entre Historia de la Ciencia y Filosofía de la Ciencia"; "Ciencia y valores éticos"; "Problemas filosóficos y metodológicos de la Economía en la Sociedad tecnológica actual"; y "La Filosofía de Imre Lakatos: Evaluación de sus propuestas".

En el presente volumen los trabajos sobre explicación científica han sido elaborados en diálogo con Wesley Salmon (Catedrático de Filosofía, U. de Pittsburgh), que fue probablemente el especialista más reconocido en este campo y que aquí aporta la versión más perfilada de dos cuestiones que estudió durante años: la diferencia entre explicaciones causales y no causales, y la estructura de la explicación causal.

Ariel Filosofía

